Molecular Simulations of Phase Behavior for Polymer Blends and Block Polymers

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Qile (Paul) Chen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

J. Ilja Siepmann, Timothy P. Lodge

May, 2018

© Qile (Paul) Chen 2018
ALL RIGHTS RESERVED

Acknowledgements

First, I would like to thank my advisors, Prof. Ilja Siepmann and Prof. Tim Lodge, for their support and guidance in the past five years. Inspired by their novel research ideas and encouraged by their hands-on teaching, I have been striving to become a scientist who is motivated by scientific curiosity and conducts research meticulously. They taught me to build my research on solid foundations and to enjoy exploring. Both the scientific knowledge and the altitude toward performing top-class research that I learned from them are invaluable to me in my future career.

I also want to thank all the people who devoted their time and efforts in helping me with my research projects. Thanks to Dr. Peng Bai and Dr. Becky Lindsey for their mentoring when I entered the group; Justine Chu and Donny Shen for the collobration on polymer projects; Rob DeJaco, Dr. Evgenii Fetisov, and Bai Xue for their help to improve the simulation code; Jingyi Chen, Bai Xue, and Dr. David Harwood for collaborations on the interfacial tension project; Dr. Mansi Shah for helping with the force field development of alkanes; Dr. Mark Schure and Dr. Kim Struksheats for getting me started on the chromatography project; Dr. Robert Hembree and Dr. Mona Minkara for many helpful discussions and critiques to my draft of manuscripts; Sujay Chopade, Yaming Jiang, Aakriti Kharel, Dr. Mohammad Savoji for helpful discussions on many projects. I want to express my gratitude to my experimental collaborators: Shuyi Xie and Prof. Reza Foudazi on the olefin/ether phase behavior project; Dr. Leonel Barreda,

Dr. Leo Oquendo, and Prof. Marc Hillmyer on the high-χ block oligomer project. I am also grateful to all Siepmann and Lodge group members, both former and present, who provided help and support to my research.

I want to thank the funding for my projects from National Science Foundation and Industrial Partnership for Research in Interfacial and Materials. Computer resources from Minnesota Supercomputing Insitute at the University of Minnesota and Argonne Leadership Computing Facility are also greatly appreciated. In addition, I want to thank Advanced Materials Technology for kindly offering a summer internship that helps the chromatography project.

I had the honor to be part of many scientific groups at University of Minnesota, including the MRSEC-IRG3, the Polymer Group, and the Chemical Theory Center. I want to acknowledge helpful discussions from students and professors in these groups.

I want to thank people who kindly mentored me during my pursue of a Ph.D.: Prof. Tom Kempf and Prof. Raul Caretta for being the instructor during my teaching assistantship; Dr. Lucas McIntosh for kindly providing helpful career advice.

Last but not least, I want to thank my family and friends, without whom a Ph.D. would have been impossible. Mom and Dad, thank you for always being by my side. I love you and want to share all my happiness and joy with you. Xinyu, special thanks to you for all your love and support through my ups and downs. Jun, Yuyang, Yiming, Lian, En, Hongyu, Tianqi, Motao, Yan, Boxin, Tuoqi, Kailong, Yaming, Ziang, Wenjia, and Zhichao, thank you for providing so much pleasure and joy to my life.

Abstract

The wide variety of phase behavior associated with polymer mixtures and block polymers enables unprecedented opportunities in developing novel polymeric materials with desired properties. However, the molecular design space of multi-component polymer systems is now so vast that guidance from theory and modeling is essential. The greatest challenge of predictive materials design is the lack of *accurate* and *precise* simulation methods in computing the phase diagram of polymer systems, due primarily to difficulties in (i) transferring polymer molecules between condensed phases and (ii) the sensitivity of phase diagram with respect to the interaction parameters used in the simulations. The overarching goal of this thesis is to address the above two problems. In this thesis, advanced sampling techniques of Monte Carlo simulations and accurate molecular models were developed to allow for the accurate and precise calculation of the mixing thermodynamics for binary mixtures. Furthermore, a case study of predictive materials design is presented, where molecular dynamics simulations were employed to explore the phase diagram of block oligomers with various chain lengths, volume fractions, and chain architectures, and thus, to guide the experimental synthesis for molecules with desired microphase morphologies. The work in this thesis lays a solid foundation for predictive materials discoveries using molecular simulations.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Polymeric materials consisting of mostly homopolymers (e.g., polyethylene, polypropylene, and polyvinylchloride) have occupied the majority of the plastics industry [1]. However, growing use of polymer blends and block polymers provides tremendous opportunities in enhancing materials properties (e.g., mechanical, optical, and electrical properties) of these commodity plastics that are beyond the reach of homopolymers [2, 3]. One key feature of such multi-component polymeric materials is that they can form a wide variety of phases (bicontinuous phase, lamallae, hexagonally-packed cylinders, etc.), with each phase having unique microstructures, and thus, exhibiting different materials properties [4]. This structure-property relationship allows for rational design of multi-component polymeric materials towards desired properties. For example, copolymers that microphase separate into percolating channels were employed in lithium ion batteries to facilitate ion transport while maintaining robust mechanical stabilities [5, 6, 7]. Similarly, use of cocontinuous polymer blends was demonstrated as an effective approach to increase the ionic conductivity and rheological properties of polymer

composites [8, 9, 10]. Non-percolating phases, such as micelles, on the other hand, are desirable for applications including plastics toughening [11, 12, 13, 14], lubricant viscosity modifiers [15], drug delivery [16, 17, 18], and filtration membranes [19, 20]. Periodic morphologies such as lamellae and hexagonally-packed cylinders formed from directed self-assembly of block polymers are under development as next-generation microelectronics and data storage devices [21, 22]. These examples have demonstrated opportunities of utilizing the diverse phase behavior of polymer blends and block polymers for developing new materials.

Despite these advantages, the use of polymer blends and block polymers also significantly increases the design space of polymeric materials, since the number of polymeric molecules consisting of certain types of repeat units increases exponentially when more components are introduced (illustrated as "exponential combinatorial explosion" by Bates et al. [23]). Therefore, predictive modeling is essential to guide the design of multi-component polymeric materials. The modeling efforts can effectively narrow down the choice of candidate molecules for target applications. In addition, molecular level insight gained from theory and modeling can greatly promote the understanding of the underlying mechanism associated with the phase behavior, and thus, elucidate the structure-property relationship.

The mainstream of thermodynamic treatments for polymer mixtures is based on the Flory–Huggins (FH) theory [24, 25]. From FH theory, a phase transition is thermodynamically driven by χ, the degree of interactions between unlike segments. While FH theory is useful in correlating experimental measurements, the real systems often exhibit deviations from the predictions of FH theory. For example, Krishnamoorti et al. [26] systematically investigated the mixing thermodynamics of binary polyolefin blends using small-angle neutron scattering. They found that about 20% of the mixtures do not conform to the FH theory, and polyolefin blends containing poly(isobutylene) even

exhibit lower critical solution temperature behavior, which is opposite to the FH theory prediction [27]. In addition, the compositional and molecular weight dependence of χ is neglected in the thermodynamic model, while many experimental [28, 29, 30, 31] and theoretical [32, 33] studies have demonstrated otherwise.

The origin of such deviation is the idealized thermodynamic treatment used in the FH theory (e.g., the mean-field assumption). These nonidealities are often lumped into the adjustable parameter in the FH theory, χ, during the correlation of experimental data. As a consequence, χ does not solely reflect the interactions between unlike segments, but it also becomes a function of liquid structure, composition, and chain length [34, 35]. This phenomenological treatment of χ makes it extremely challenging to predict χ from molecular structures, and thus, significantly undermines the predictive utility of the FH theory.

Such deficiencies of the existing theories motivate efforts on the theoretical front to quantitatively predict how molecular level interactions influence mixing thermodynamics. Graessley and co-workers used equation-of-state modeling to correlate the mixing behavior of polyolefin blends from small-angle neutron scattering [36, 26]. Bates et al. incorporated the mismatch in persistence lengths of two components into the FH theory [37, 38]. Similarly, both the polymer reference interaction site model (PRISM) developed by Schweizer and co-workers [39, 40] and the lattice cluster theory by Freed et al. [41, 42] considered the difference in local packing for two components in their models. Lipson et al. used locally correlated lattice models and concluded that mismatches in cohesive energy densities and free volumes can be used to compute χ [43]. Despite this progress, one key unsolved challenge is that many parameters used in the above theories are obtained from experimental measurements [43] rather than from first-principle calculations. This drawback limits these theories from exploring the phase behavior that

involves new polymer molecules that have not yet been synthesized, and thus, from addressing the "exponential combinatorial explosion" challenge for the development of new polymeric materials.

Molecular simulations have much to offer in providing predictive capabilities for polymeric materials design. Monte Carlo (MC) and molecular dynamics (MD) simulations, two major categories of simulation techniques, are extensively used to sample the configurational phase space of polymer systems and compute their structural, thermodynamic, and transport properties [44]. In the simulations, one or several atoms are represented by one pseudo-atom, and the interactions between pseudo-atoms (i.e., the force field) is "calibrated" to either quantum mechanical or experimental data. The transferability of the force field parameters minimizes the size of the training set and guanrantees the predictive capability of the simulations. Furthermore, molecular simulations can also access molecular level information that is difficult to extract from experiments, such as the chain conformation of polymers and the local packing structure. These data can provide valuable insight into the mechanistic investigation of the phase behavior for polymeric mixtures and block polymers.

Despite its potential usefulness, the shadow over the predictive design of polymeric materials using molecular simulations is the lack of *accurate* and *precise* simulation methods as well as *accurate* molecular models in computing the phase diagram of multicomponent polymer systems. The first challenge is to transfer polymeric molecules between condensed phases, which is needed in computing the chemical potential of polymers or directly calculating the phase diagram [45]. The acceptance of such molecular transfer is nearly zero for long-chain molecules using conventionally insertion methods [45]. Therefore, many simulation studies used indirect methods to compute χ, which relied on theoretical frameworks to interpret the simulation data. For example, a method analogous to neutron scattering experiment was developed [46, 47], where the structure

factor is computed from MD simulations of miscible binary blends, and then, the random phase approximation (RPA) [48] is applied to extract χ. This method was extensively used to calculate χ for polyolefin mixtures [49, 50, 51]. However, the main shortcoming of such method is that the accuracy of the theoretical framework (RPA in this case) is still a matter of discussion. For instance, Miquelard-Garnier and Roland summarized the literature data and found some inconsistencies between χ deduced from RPA and other experimental methods [52].

This barrier prompted numerous attempts to develop enhanced sampling techniques to enable the transfer of polymeric molecules between condensed phases, and thus, the *accurate* and *precise* calculation of χ. One idea is to bias the selection of candidate sites during the chain growth towards more energetically favorable positions to increase the acceptance of molecular transfer [53, 54, 55, 56]. In addition, intermediates or "fractional components" of the polymer chain are used to divide the insertion of the entire polymer molecule into several stages and compute the incremental chemical potentials [57, 58, 59, 60, 61, 62, 63]. Other simulation methods such as the connectivity-altering MC move [64, 65] allow for drastic conformational changes of polymer chains, and thus, significantly increase the sampling efficiency. Recently, Zhang et al. developed a thermodynamic integration method to compute χ on the path along which one polymer transforms into the other [63], which shed light on the potential of computing χ free from the RPA.

The second difficulty associated with predictive modeling of polymer blends and block polymers is the *accuracy* of the molecular models. For atomistic level simulations, the phase diagram of polymer systems (in fact, liquid–liquid phase diagram in general) is found to be extremely sensitive to force field parameters. A recent simulation of the alkane/ethanol mixture phase diagram [62] shows that a difference of 1 kJ/mol in the free energy of mixing, often considered as a "small" deviation in the force field parameterization, results in about 50 K overestimation in the critical temperature of the binary

liquid–liquid phase diagram. The challenge is heightened by the fact that coarse-graining is sometimes essential to model high molecular weight polymers [66, 67, 68, 69], which is both a blessing and a curse. On the one hand, the coarse-graining permits modeling of high molecular weight polymers that is otherwise too computationally expensive using atomistic simulations. Papakonstantopoulos et al. [70] and Chremos et al. [71] studied the self-assembly of poly(styrene-*b*-methyl methacrylate) using coarse-grained molecular models that represent five to six repeat units by one pseudo-atom. On the other hand, however, the interactions and structures on the atomistic level are often obscured in the coarse-grained models. Unfortunately, atomistic level interactions can remarkably affect the accuracy of the prediction, as it is known that χ is sensitive to these interactions.

In this thesis, the goal is to address the above two challenges by using molecular simulations with advanced sampling techniques and accurate force fields. I aim to develop state-of-the-art computational methods and molecular models to *accurately* and *precisely* calculate χ and the associated phase behavior of polymer mixtures and block polymers. Such advances can be applied to achieve the *in silico* design of polymeric materials with desired properties and guide the experimental synthesis. In addition, the structural information from these simulations can significantly enhance the fundamental understanding of the phase behavior for polymer mixtures and block polymers.

1.2 Thesis Overview

The overarching goal of this thesis is to develop robust computational methods in predicting the phase behavior of polymer blends and block polymers from their molecular structures. The simulation methods including Gibbs ensemble Monte Carlo (GEMC) simulations and MD simulations were employed to study a variety of polymer blends and block polymers.

Chapter 2 discusses the development of the simulation method to compute the thermodynamic and structural properties associated with the miscibility of binary olefin oligomer mixtures. GEMC simulations were used to study oligomers representing poly(ethylene-*alt*-propylene), polypropylene and head-to-head polypropylene. The cohesive energy densities, Π_{CED}, of different oligomers that are used in estimating the χ of binary mixtures but are not measurable for polymers, were calculated from simulations. The extrapolation of simulation data for C5 to C36 oligomers enables the determination of the Π_{CED} values of infinite-chain-length polyolefins. The results show remarkably good agreement with values deduced from neutron scattering experiments on binary polyolefin mixtures. Furthermore, the Flory-Huggins χ parameters were computed from the Monte Carlo simulations free from the RPA. The stabilized irregular mixing behavior for the binary propylene and head-to-head propylene mixtures is consistent with their polymeric counterparts. The chain-length insensitivity of the mixing thermodynamics is rationalized using insights from the structural analysis. Results in this chapter identified GEMC simulations as a promising route to predict and understand the phase behavior of polymer mixtures.

Chapter 3 documents results from molecular simulations to understand the molecular weight dependence of the Flory-Huggins χ parameter and the effect of dispersity on the phase diagram of binary polymer mixtures. The binary mixtures containing oligomeric poly(ethylene-*alt*-propylene) (PEP) and poly(ethylene oxide) dimethyl ether (PEO) were used as the model system. First, χ for PEP/PEO mixtures with various molecular weights was computed from GEMC simulations. An empirical model with an adjustable parameter was developed to rationalize the dependence and enable the accurate prediction of χ for PEP/PEO mixtures with arbitrary molecular weights. Second, the effects of dispersity of PEO on the binary PEP/PEO phase diagram were probed

from simulations of liquid–liquid phase equilibria. Coexistence curves for mixtures containing PEO with various molecular weight distributions and dispersities but with the same number average molecular weight were compared and discussed. In addition, the aggregation behavior and microscopic heterogeneity of the polymer mixture was probed from structural analysis of simulation trajectories. The results in this chapter allow for the accurate prediction of (i) χ for polymer mixtures with arbitrary molecular weights and (ii) the liquid–liquid phase diagram of polydisperse binary polymeric mixtures from MC simulations.

Chapter 4 presents results from molecular simulations that are used to predictively design a series of high-χ block oligomers (HCBOs) for self-assembly into microphases with domain sizes that are as small as 1 nm. MD simulations with atomistic models were used to explore the HCBOs with a wide range of chain lengths N, volume fractions f, and chain branching at three temperatures. Ordered microphases that include lamellae, perforated lamellae, and hexagonally-packed cylinders were observed from these simulations. The smallest domain periods are 3.0 nm and 2.1 nm for lamellae and cylinders, respectively, and the smallest size of the polar domains is approximately 1 nm. It is intriguing that the detailed phase behavior of HCBOs is different from that of their polymeric counterparts. It is found that the complex phase behavior of HCBOs is a product of interplay between both "surfactant factors" (such as the head group interactions, chain flexibility, and interfacial curvature) and "block polymer factors" (including χ, N, and f). The hydrogen bonding and structural analyses from simulation trajectories facilitate the understanding of the phase behavior for HCBOs at the sub-5 nm length scale, which can potentially aid the design of HCBOs tailored towards particular desired morphologies.

Chapter 5 describes that use of k-d tree, a data structure to store the coordinates of the particles in a simulation, can substantially accelerate Monte Carlo simulations for a

wide range of molecular systems, including polymers. The efficiency enhancements due to the k-d tree data structure were assessed from MC simulations for various molecules in different ensembles. A rule of thumb was generalized as to which systems benefit most from the use of k-d tree, and the rationale for the speed-up was investigated. The increase in simulation efficiency can boost the throughput of simulation studies and enable the drive towards the computational screening for materials discovery.

In summary, the work in this thesis has demonstrated that molecular simulation is a powerful tool in the fast, accurate, and precise determination of the phase diagram of polymer blends and block polymers, which opens up a new avenue for predictive design of polymeric materials with tailored properties.

Chapter 2

Molecular Simulation of Olefin Oligomer Blend Phase Behavior

2.1 Introduction

Saturated hydrocarbon polymers, known as polyolefins, are used in materials that are of significant commercial importance [1]. Polyolefin blends, similar to alloys, give rise to novel morphologies and material properties that are often not accessible with pure polyolefins [2]. These interesting structures and properties are often dictated by the segment–segment interactions, characterized by the binary interaction parameter, χ. For a binary mixture, χ is defined in the Flory–Huggins (FH) theory as:

$$\frac{\Delta G_{\mathrm{m}}}{k_{\mathrm{B}}T} = \frac{\phi_1}{N_1}\ln(\phi_1) + \frac{1-\phi_1}{N_2}\ln(1-\phi_1) + \chi(T)\,\phi_1(1-\phi_1) \tag{2.1}$$

This chapter is reproduced in part with permission from (Q. P. Chen, J. D. Chu, R. F. DeJaco, T. P. Lodge and J. I. Siepmann, "Molecular Simulation of Olefin Oligomer Blend Phase Behavior", *Macromolecules* **2016**, *49*, 3975–3985, American Chemical Society)

where ΔG_{m}, k_{B}, and T are the free energy of mixing per reference site, the Boltzmann constant, and the absolute temperature, respectively; ϕ_1 and N_i ($i = 1$ or 2) are the volume fraction of component 1 and the ratio between molar volume of component i and the reference site molar volume, v_0.

The correlation between the phase behavior of a binary mixture and the associated χ parameter is relatively well understood, but accurate predictions of the χ parameter from knowledge of molecular structure alone are still lacking. These predictions are essential in understanding the structure–property relationship of polymers and for aiding the design of next-generation polymeric materials. The Hildebrand formalism provides the relationship between χ and the solubility parameters, δ_i, of the two components [72]:

$$\chi_{\mathrm{SP}}(T) = \frac{v_0}{k_{\mathrm{B}}T}\Big[\delta_1(T) - \delta_2(T)\Big]^2 \tag{2.2}$$

However, the χ parameter from eq 2.2 may not be sufficient to accurately describe the phase behavior of some binary polyolefin mixtures. As an aside, it should be noted that shortcomings of the Hildebrand formalism are more prevalent when moving beyond mixtures of saturated hydrocarbons [73, 74]. Extensive small-angle neutron scattering (SANS) studies for polyolefin blends [75, 76, 77, 36, 26, 78], where χ can be fitted from the scattering pattern with the random phase approximation [48], quantified the deviation of the χ_{SANS} parameter from the prediction of eq 2.2 by the mixing irregularity, X_{E}, as follows

$$X_{\mathrm{E}} = \frac{k_{\mathrm{B}}T}{v_0}\chi_{\mathrm{SANS}} - \Big[\delta_1(T) - \delta_2(T)\Big]^2 \tag{2.3}$$

Some combinations of polyolefins yield X_{E} with an absolute value greater than 2.5×10^{-3} MPa, which is classified as irregular mixing behavior [75, 76, 77, 36, 26, 78].

A better understanding of this anomalous behavior can lead to more accurate predictions of χ parameters for blends, but the origin of the irregular mixing behavior

is still a matter of controversy. A wealth of experimental, simulation, and theoretical studies have supported various hypotheses regarding mixing irregularities. Graessley et al. [36, 26] interpreted this deviation as a departure of the cross-component interactions from the Berthelot combining rule, that is often also invoked for equation-of-state modeling. Bates and co-workers [37, 38] suggested a correlation between X_{E} and the mismatch in the persistence lengths of two components, which implies that the irregular mixing is mainly entropic in origin. Similarly, simulation studies by Maranas et al. [35, 79, 80] indicated a correlation of X_{E} with the shapes of single-component radial distribution functions (RDFs), as well as with the molecular packing pattern. On the other hand, Schweizer and co-workers [81, 82, 39, 40, 83, 84] suggested the coupling of cohesive energy and local packing in their theoretical studies, attributing the irregularity to an enthalpic contribution. Likewise, Wu and co-workers [49, 50] pointed to the role of the intramolecular energy in the miscibility of polyolefins. From lattice cluster theory studies, Freed and co-workers [41, 85, 42] concluded that the mixing anomaly can be either entropic or enthalpic depending on the system. This is also supported by more recent studies by Lipson et al. [43, 86, 87, 88] using locally correlated lattice models that indicate a mismatch in either free volume or cohesive energy density resulting in irregular mixing behavior.

Molecular simulations using transferable force fields to represent the interactions of the olefin isomers have much to offer for elucidating the molecular-level origin of this anomalous mixing behavior. The first problem towards the quantitative prediction of mixing irregularities is the accurate and precise calculation of Π_{CED} or δ of a polymer given by

$$\delta(T,p) = \sqrt{\Pi_{\mathrm{CED}}(T,p)} = \sqrt{\frac{\Delta U_{\mathrm{vap}}(T,p)}{V_{\mathrm{liq}}(T,p)}} = \sqrt{\frac{\Delta H_{\mathrm{vap}}(T) - RT}{V_{\mathrm{liq}}(T)}} \tag{2.4}$$

where ΔU_{vap}, ΔH_{vap}, and V_{liq} are the molar internal energy change upon moving a particle from the condensed liquid phase to being an isolated molecule, the molar heat of vaporization, and the liquid-phase molar volume, respectively. The last equality is only valid at the saturated vapor pressure (and for a vapor phase obeying the ideal gas law), but this is the only state point where these values can be determined directly from experimental measurements. Unfortunately, Π_{CED} (and hence also δ) cannot be determined for polyolefins because their exceedingly low saturated vapor pressures at temperatures where decomposition does not occur, makes measurement of ΔH_{vap} impossible. As a consequence, a common approach is to substitute the internal pressure, Π_{IP}, for Π_{CED} because the former is available from equation-of-state PVT measurements [26]. However, it was shown previously by experiment, simulation, and theory that this substitution is neither accurate nor precise [26, 34, 89, 43]. In contrast, simulations can provide a more reliable way to estimate polymeric Π_{CED} values. Maranas et al. [34] simulated a series of olefin oligomers with 7 to 30 carbon atoms, and the infinite-chain-length Π_{CED} was extrapolated from these oligomer data. Guenza et al. [51, 90] adopted a different approach, where a coarse-graining scheme was developed to compute Π_{CED} of polyolefins that have 96 carbons.

The second problem is a lack of simulation methods to calculate directly the χ parameter. In previous work, the calculation of χ was performed by either neglecting the entropic contribution to the excess free energy [91, 49], or by fitting χ_{SANS} from the scattering pattern calculated from large-scale molecular dynamics simulations [46, 49, 50, 51, 90]. The former method only works for regular solutions, and thus may not be valid for polyolefin mixtures. For the latter approach, a large system size is essential, which makes the simulations very expensive, and assumptions are involved in the analysis of the scattering signal. Gibbs ensemble Monte Carlo (GEMC) [92, 93] is the method of choice for the direct calculation of phase equilibria of molecular systems, but the low

acceptance in transferring long chain molecules between two phases has impeded its usage in polymer simulations [94].

In this work, GEMC simulations on olefin oligomers are performed with a variety of simulation techniques to mitigate the difficulties mentioned above. Three series of oligomers are studied in this work, namely, poly(ethylene-*alt*-propylene) (PEP), atactic polypropylene (PP), and head-to-head polypropylene (hhPP). Their repeat unit structures are shown in Figure 2.1. Simulations focused on oligomers with the number of carbon atoms, N_C, ranging from 5 to 36. Here, these molecules will be named as (polymer name)-C(number of carbons). For example, hhPP-C18 represents the molecule that shares the same repeat unit structure as hhPP and contains 18 carbon atoms (and 38 hydrogen atoms). This chapter addresses two topics: First, results for unary systems are presented, with the emphasis being placed on the disparity between Π_{CED} and Π_{IP}, as well as a more accurate approach for estimating Π_{CED} and δ for polymers from these oligomer simulations. Second, the calculation of χ for binary olefin oligomer mixtures is described. The results show that similar mixing irregularities are present for oligomeric mixtures as for their polymeric counterparts. Structural analysis provides some insights into this similarity and the origin of the irregular mixing behavior.

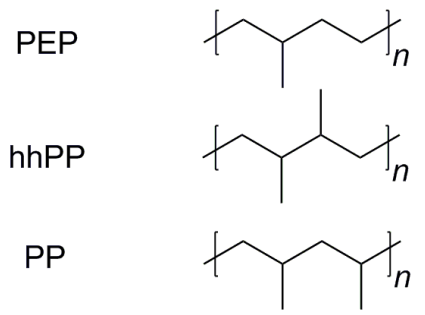

Figure 2.1: Repeat unit structures of the olefin oligomers studied in this work.

2.2 Simulation Details

The united-atom version of the TraPPE (transferable potentials for phase equilibria) force field was used in this work [95, 96]. These models are semiflexible with fixed bond length, but bending and dihedral angles are allowed to change. The non-bonded interactions of CH_x groups (with $x = 1$, 2, or 3) were described by Lennard-Jones (LJ) potentials. Lorentz–Berthelot combining rules [97] were used to determine the LJ parameters for all unlike interactions. A spherical truncation at 14 Å was used, and analytical tail corrections were applied to account for the interactions beyond the truncation distance [98]. The in-house MCCCS–MN (Monte Carlo for Complex Chemical Systems–Minnesota) software program was used for all simulations [99].

GEMC simulations [92, 93] in the canonical (NVT) ensemble were used to simulate vapor–liquid equilibria of pure oligomers. The system sizes are as follows: 800 molecules for C5 and C6; 400 molecules for C10, C12, and C15; 300 molecules for C18, C20, C24, and C25; and 200 molecules for C30 and C36. The simulations were initialized with all the molecules being placed in the liquid phase. Six types of MC moves were applied, including translational, rotational, conformational, volume exchange, particle transfer, and inter-box identity switch moves [100, 101, 102]. The coupled–decoupled configurational-bias Monte Carlo (CD-CBMC) algorithm [96] was used to enhance the acceptance rates of the conformational, particle transfer, and identity switch moves. It should be noted that the present implementation of CD-CBMC does not include contributions from improper torsions that are required to preserve tacticity [49]. Improper torsions can be added to allow for the investigation of oligomers with specific tacticity [103, 104] in future work. The attempted moves were distributed as follows: 4.5% translational, 4.5% rotational, 40% conformational, 1% volume exchange, 10% swap, and 40% identity switch moves. Reasonable variations in move ratio settings were tested, but no significant discrepancies in results and uncertainties were identified.

To overcome the low acceptance probabilities of direct particle transfer moves in GEMC simulations, indirect particle transfers were performed by the use of impurity molecules and identity switch moves [100, 101, 102]. There are two molecules of each impurity type, that are distributed on average evenly between the two phases with the aid of biasing potentials. Impurity molecules share the same repeat unit structure as the target molecule; for example, the simulation of hhPP-C24 includes hhPP-C6, hhPP-C12, and hhPP-C18 as impurities (resulting in a dispersity of 1.006). Instead of the direct swap of an hhPP-C24 molecule, identity switch moves were performed on neighboring chain length pairs, e.g., C6 with C12, C12 with C18, and C18 with C24. Each identity switch pair was chosen with equal probability, and swap moves were only carried out for the shortest impurity molecule, e.g., hhPP-C6. This approach divides the transfer of molecules into several stages, which reduces the number of interaction sites to grow in one stage, and thus boosts the acceptance significantly [102, 105]. The accuracy of the simulations can be retained via dispersity corrections which assume that extensive properties are linear combinations of contributions from all species, i.e., the ΔH_{vap} reported is only that for the compound of interest. For validation of this approach, GEMC simulations with and without impurities were performed at $T = 590$ K for hhPP-C24 and PP-C24 (where direct particle transfer moves have sufficiently high acceptance rates), and statistically equivalent results were obtained for all properties of interest.

All GEMC simulations were equilibrated until there was no drift in both energies and number densities for the two phases. Based on these criteria, the equilibration period was found to require from 2.5×10^5 Monte Carlo cycles (MCCs, consisting of N randomly selected moves, where N is the number of molecules in the system) for C12 to 10^6 MCCs for C36. The production periods encompass 2.4×10^6 and 1.5×10^6 MCCs for the C12 and C24 oligomers of PP and hhPP in order to reduce the statistical uncertainties for

the χ calculation, while for other chain lengths, only 3×10^5 MCCs were used for the production runs.

Single-molecule simulations were conducted in order to obtain the vapor-phase internal energy of the olefin oligomers. In these simulations, the periodic boundary condition was not applied. Only CBMC moves were used, and the simulations consisted of 10^6 MC moves. Simulations in the NpT ensemble (with the same number of molecules as for the GEMC simulations) were used to determine the bulk liquid-phase properties of olefin oligomers at $p = 1$ atm. In addition to translational, rotational, volume and regular CBMC moves, self-adapting fixed-end-point CBMC (SAFE-CBMC) approach [106] was applied to ensure good sampling of the internal conformation of long chain molecules. A 5 Å cutoff was used to compute Rosenbluth weights during the CD-CBMC and SAFE-CBMC growth process, and then corrected to the 14 Å cutoff in the acceptance rule [107]. MC moves were divided as 10% translational, 10% rotational, 1% volume, 55% regular CBMC, and 24% SAFE-CBMC moves. These NpT simulations were equilibrated for 2.5×10^5 MCCs and run for another 5×10^5 MCCs as the production period.

The values of Π_{CED} and δ were calculated via the following equation:

$$\delta = \sqrt{\Pi_{\mathrm{CED}}} = \sqrt{\frac{U_{\mathrm{iso}} - U_{\mathrm{liq}}}{V_{\mathrm{liq}}}} \tag{2.5}$$

where U_{iso} is the molar internal energy of the isolated molecule from the single-molecule simulations; U_{liq} is the molar internal energy of the pure melt. Both U_{liq} and V_{liq} were computed from either unary NVT-GEMC simulations or NpT ensemble simulations (only when $p_{\mathrm{sat}} \leq 1$ atm), and no statistically significant differences were observed between the two approaches. The thermal expansivity, α, and isothermal compressibility, β, were computed from fluctuations in the NpT ensemble simulations [108, 109]. Π_{IP}

can then be computed as follows

$$\delta_{\mathrm{IP}} \equiv \sqrt{\Pi_{\mathrm{IP}}} = \sqrt{\frac{T\alpha}{\beta}}. \tag{2.6}$$

where δ_{IP} denotes the approximated solubility parameter from Π_{IP}.

Binary olefin oligomer mixture simulations were set up in a similar way as those for unary systems, and either NVT or NpT Gibbs ensemble simulations are used. MC moves were divided between different components based on the total number of degrees of freedom that each component possesses. The Gibbs free energy of mixing, ΔG_{m}, per molecule was computed from vapor phase number densities by neglecting the pressure effect on mixing:

$$\frac{\Delta G_{\mathrm{m}}}{k_{\mathrm{B}}T} = x_1 \ln\left(\frac{\rho_1^{\mathrm{mix}}}{\rho_1^{\mathrm{pure}}}\right) + (1-x_1)\ln\left(\frac{\rho_2^{\mathrm{mix}}}{\rho_2^{\mathrm{pure}}}\right) \tag{2.7}$$

where x_1 is the mole fraction of component 1 in the liquid phase; ρ_1^{mix}, ρ_1^{pure}, ρ_2^{mix}, and ρ_2^{pure} are the number densities of component 1 or 2 in the pure melt or mixture vapor phases, respectively. The FH χ parameter was calculated from the excess Gibbs free energy (excess chemical potential) of mixing, $\Delta G_{\mathrm{m}}^{\mathrm{E}}$, per molecule:

$$\chi_{\mathrm{CP}} = \frac{\Delta G_{\mathrm{m}}^{\mathrm{E}}}{k_{\mathrm{B}}T\phi_1(1-\phi_1)N} \tag{2.8}$$

where N is the number of repeat units. The mixing irregularity as well as its enthalpic and entropic parts were calculated as follows:

$$X_{\mathrm{E}} = \frac{k_{\mathrm{B}}T}{v_0}\chi - (\delta_1 - \delta_2)^2 \tag{2.9}$$

$$X_{\mathrm{E}}^{\mathrm{H}} = \frac{k_{\mathrm{B}}T}{v_0}\chi^{\mathrm{H}} - (\delta_1 - \delta_2)^2 = \frac{\Delta H_{\mathrm{m}}}{v_0\phi_1(1-\phi_1)N} - (\delta_1 - \delta_2)^2 \tag{2.10}$$

$$X_{\mathrm{E}}^{\mathrm{S}} = \frac{k_{\mathrm{B}}T}{v_0}\chi^{\mathrm{S}} = \frac{-T\Delta S_{\mathrm{m}}^{\mathrm{E}}}{v_0\phi_1(1-\phi_1)N} \tag{2.11}$$

where the enthalpy of mixing, ΔH_{m}, was calculated from the difference in the molar volumes and internal energies upon mixing, while the excess entropy of mixing, $\Delta S_{\mathrm{m}}^{\mathrm{E}}$, was calculated from the difference of ΔH_{m} and $\Delta G_{\mathrm{m}}^{\mathrm{E}}$.

For all numerical data, the statistical uncertainties were computed from either eight or sixteen independent simulations and are reported as 95% confidence interval in this chapter.

2.3 Results and Discussion

2.3.1 Single-component thermodynamic properties

Liquid Density and Vapor Pressure. The simulation data for the liquid densities of oligomers are displayed in Figure 2.2. Results from NVT-GEMC and NpT ensemble simulations are indistinguishable, despite the small difference in the pressure of these simulations. Experimental data are only available for low-molecular-weight compounds and their polymeric counterparts, but not for the oligomers. For species with $N_{\mathrm{C}} = 5$ and 6, the liquid densities yield unsigned percentage errors smaller than 1% for all three oligoolefins at three temperatures compared with the experimental data [110]. For $N_{\mathrm{C}} = \infty$, the following empirical equation was used to fit simulation data and extrapolate to the densities of high-molecular-weight polymers:

$$\frac{1}{\rho} = a + \frac{b}{N_{\mathrm{C}}} + \frac{c}{(N_{\mathrm{C}})^2}, \tag{2.12}$$

where a, b and c are fitting parameters. A similar functional form but without the quadratic term is known to fit the experimental data of n-alkanes well [111], which implies that the reciprocal of the density, or the mass-specific volume, can be divided into an infinite-chain-length contribution and a chain-end correction. Here, a quadratic term is added to account for the non-linearity of the chain-end effect due to the presence

of the branches. Good agreement between extrapolated densities and corresponding experimental values [36, 26] is observed (see Table 2.1). Note that PEP-C5 was not included in the fitting at $T = 440$ K, because it is too close to the critical temperature [112] ($T_c = 461$ K).

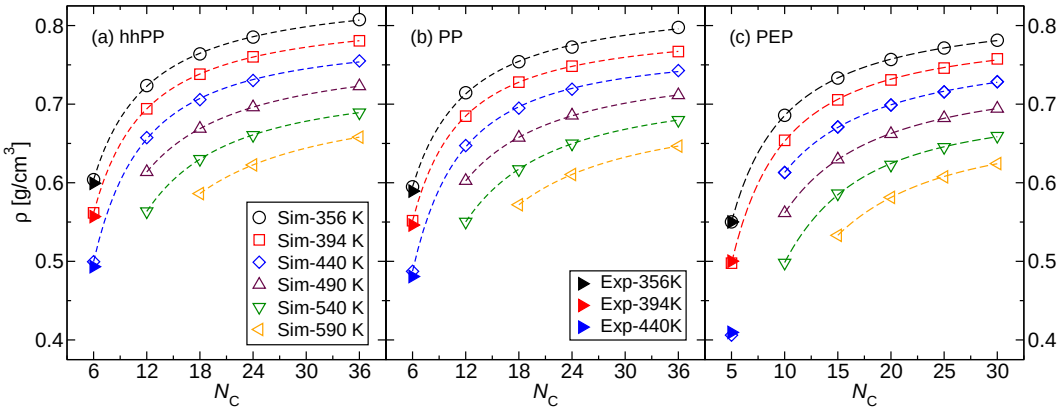

Figure 2.2: Liquid densities of hhPP, PP, and PEP oligomers. Line styles and symbols are the same for all three parts. The experimental data are from the compilation of Smith and Srivastava [110]. Dashed lines represent the fits of simulation data to eq 2.12. Statistical uncertainties are not shown because they are smaller than the size of the symbols.

The saturated vapor pressures of the olefin oligomers are shown in Figure 2.3. Compared to the experimental data [110] for $N_C = 5$ and 6, the simulation results overestimate the vapor pressures for all three molecules by around 10 to 40% (with the deviations decreasing as the temperature increases). This overestimation of vapor pressures has previously been reported for alkanes modeled by the TraPPE united atom force field [95, 96], and more expensive explicit-hydrogen models yield more accurate vapor pressures (and also heats of vaporization) [113].

Solubility Parameter and Cohesive Energy Density. Figure 2.4 shows Π_{CED} of olefin oligomers as function of inverse chain length. As also observed for the liquid

Table 2.1: Extrapolated Infinite-Chain-Length Densities[a] from Oligomer Simulations and the Corresponding Percentage Error ($\Delta\rho$) Compared with Experimental Polymer Data [36, 26].

	hhPP		PP		PEP	
T [K]	ρ [g/cm^3]	$\Delta\rho$ [%]	ρ [g/cm^3]	$\Delta\rho$ [%]	ρ [g/cm^3]	$\Delta\rho$ [%]
356	0.849_3	1.2_4	0.830_3	0.4_4	0.827_2	0.8_3
394	0.819_3	0.2_4	0.802_3	-0.2_4	0.798_1	-0.3_2
440	0.790_1	-0.2_2	0.773_2	-0.5_3	0.782_2	0.7_3
490	0.775_3	n/a	0.763_3	n/a	0.746_3	n/a
540	0.740_3	n/a	0.731_4	n/a	0.698_3	n/a
590	0.719_8	n/a	0.708_8	n/a	0.687_6	n/a

[a]The statistical uncertainties in the last digit are given as the subscripts.

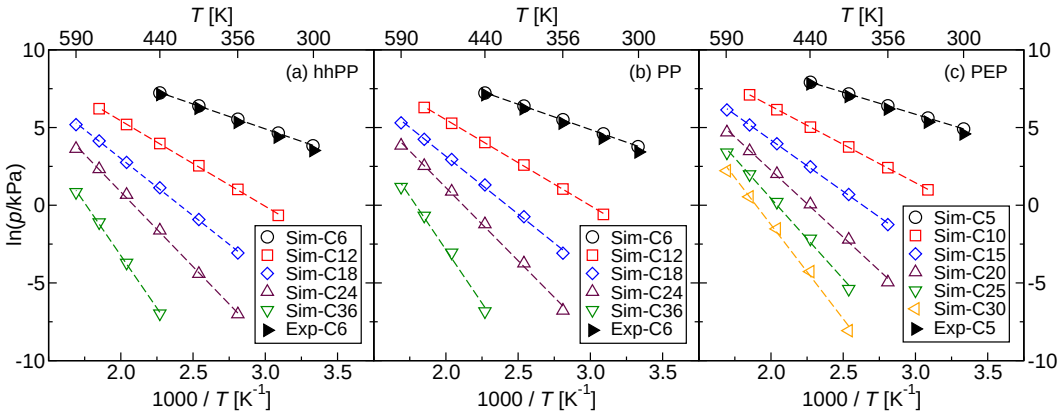

Figure 2.3: Saturated vapor pressures of hhPP, PP, and PEP oligomers. The experimental data are taken from Smith and Srivastava [110]. The dashed lines are linear fits of the simulation data.

densities, NVT-GEMC and NpT ensemble simulations yield statistically indistinguishable values for Π_{CED} and δ. Compared to the experimental data [110], the simulations for chains with $N_{\mathrm{C}} = 5$ and 6 underestimate Π_{CED} by around 5% with the deviations being larger at lower temperatures. This is due to the underestimation of ΔH_{vap} for

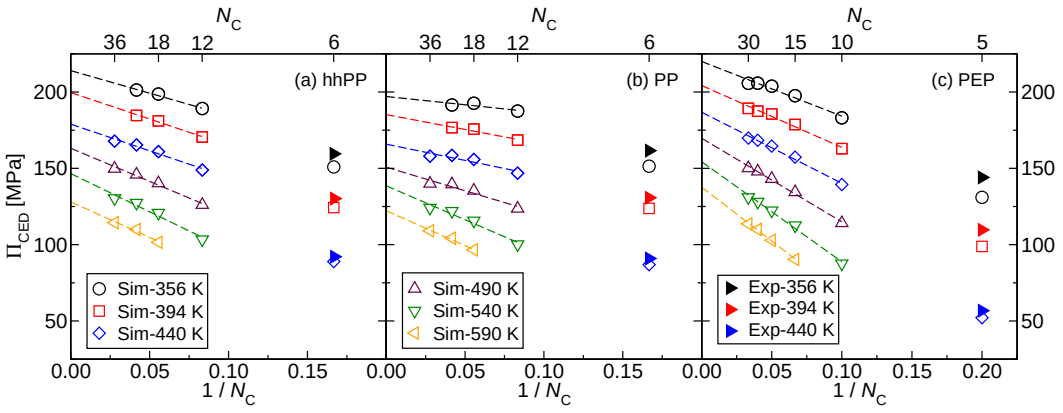

Figure 2.4: Π_{CED} values of hhPP, PP, and PEP oligomers. Line styles and symbols are the same for all three parts. The experimental data are taken from Smith and Srivastava [110]. The dashed lines are weighted linear fits of simulation data with $N_C \geq 10$.

united-atom models [95, 96], and is consistent with the vapor pressure data that indicate that those deviations decrease with increasing temperature. In other words, the slope of the Clausius–Clapeyron plot, which is approximately proportional to the heat of vaporization, is underestimated for these models.

Similar to the simulation studies by Maranas et al. [34], a linear relation between Π_{CED} and $1/N_C$ is found for $N_C \geq 10$. In some cases, the Π_{CED} values for the compounds with the largest N_C at a given temperature fall somewhat below the linear fit, but these data also have the largest statistical uncertainties. From this linear relationship, infinite-chain-length Π_{CED} and solubility parameters are obtained and displayed as a function of temperature in Figure 2.5. For all three species in this study, δ values are found to vary linearly with temperature, and linear fits yield

$$\delta_{\mathrm{hhPP}}(T)/\mathrm{MPa}^{1/2} = -0.0140(T/\mathrm{K}) + 19.63 \tag{2.13}$$

$$\delta_{\mathrm{PP}}(T)/\mathrm{MPa}^{1/2} = -0.0129(T/\mathrm{K}) + 18.69 \tag{2.14}$$

$$\delta_{\mathrm{PEP}}(T)/\mathrm{MPa}^{1/2} = -0.0131(T/\mathrm{K}) + 19.47 \tag{2.15}$$

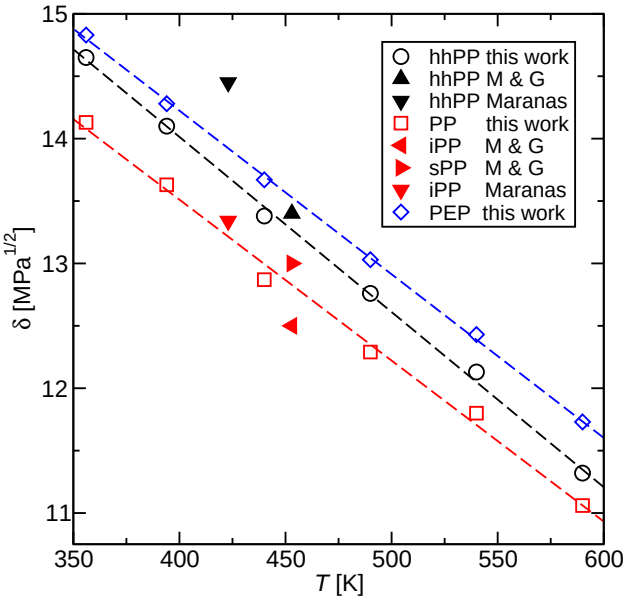

Figure 2.5: Temperature dependence of inifinite-chain-length δ values for hhPP, PP, and PEP polymers obtained from simulation studies. The dashed lines are fits to the δ values obtained in this work (open symbols; statistical errors are smaller than the symbol size). The filled black and red down triangles represent data for hhPP and isotactic PP from Maranas et al. [34]. The filled black up, red left, and red right triangles represent data for hhPP, isotatic PP, and syndiotactic PP from McCarty and Guenza [51].

These empirical relationships allow the estimation of δ at any temperature within the range of this study. In addition, these extrapolated polymeric δ values are in good agreement with results from other simulation studies [34, 51] (with the hhPP value reported by Maranas et al. [34] being the only outlier).

Figure 2.6 shows a comparison between simulation and SANS experiments [36, 78] for the difference in solubility parameters of the three pairs of olefins investigated here. The good match between simulation and experimental SANS data implies that accurate Π_{CED} and δ values for polymers can be obtained via the extrapolation of oligomer simulation data. In contrast, if Π_{CED} is approximated by Π_{IP}, then the resulting δ

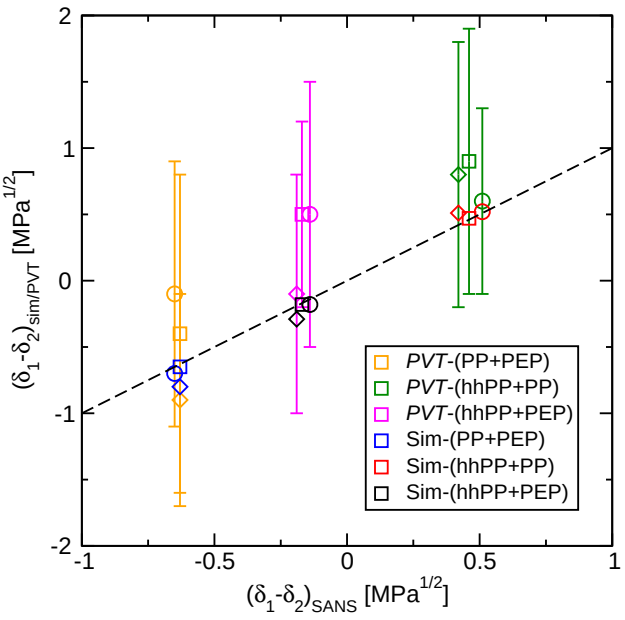

Figure 2.6: Differences in infinite-chain-length solubility parameter of three binary pairs at at $T = 356$ K (circles), 394 K (squares), and 440 K (diamonds). Data are shown for predictions from simulation and experimental PVT measurements [26] (where δ_{IP} is measured instead of δ) versus data from SANS experiments [36, 78].

differences do not agree well with the SANS data, and suffer from uncertainties that are larger than 0.5 MPa$^{1/2}$. This result implies that the χ parameter computed from eq 2.2 will also deviate from its true value when Π_{IP} is used as substitute for Π_{CED}. The discrepancy between Π_{CED} and Π_{IP} values will be discussed further below. In summary, the linear dependence of Π_{CED} on $1/N_C$ already holds for oligomers with $N_C \geq 10$ and can be exploited to reduce the chain length (and hence computational cost) required to predict Π_{CED} for polymers. However, judicious choices of the lower and upper bounds of the oligomer length range will be needed to take significantly different persistence lengths into account.

Comparison of Π_{CED} and Π_{IP}. As mentioned earlier, Π_{IP} is widely used to approximate Π_{CED} for polyolefins. However, the validity of this approach is in doubt. The ratio of $\mu = \Pi_{\mathrm{CED}}/\Pi_{\mathrm{IP}}$ is found to consistently fall below unity for polyolefins when different approaches are combined to estimate both Π_{CED} and Π_{IP} [26, 35, 43]. In this work, both quantities are calculated for PEP-C5, PEP-C10 and PEP-C30 at $T = 394$ K, and results are shown in Figure 2.7. Comparison of Π_{CED} and Π_{IP} at $N_{\mathrm{C}} = 5$, 10 and 30 shows that μ is smaller than unity for all three chain lengths. Moreover, μ appears to decrease monotonically as the chain length increases. From extrapolation, μ at infinite chain length is determined to be 0.74 ± 0.08, that is in excellent agreement with the experimental value of 0.72 ± 0.11 [26], and another simulation prediction of 0.75 ± 0.02 [34]. Here it should be noted that extrapolation of the linear fits to N_{C} would point to an intersection point but, as discussed previously, Π_{CED} for $N_{\mathrm{C}} = 5$ falls well below the linear fit because ΔH_{vap} falls off rapidly as the critical temperature of a compound is approached. In contrast, the Π_{IP} value for $N_{\mathrm{C}} = 5$ is close to the extrapolation from the linear fit because non-linear effects in thermal expansivity and isothermal compressibility partially offset each other.

Π_{CED} and Π_{IP} values for hhPP, PP, and PEP oligomers at three different temperatures are compared in Figure 2.8. In general, the disparity persists and all μ values fall below unity. Results show that μ depends on repeat unit structure, chain length, and temperature. Among them, the chain length has the largest effect on the value of μ when data for $N_{\mathrm{C}} < 10$ are included. Given that all these factors play a role, it is evident that a common value of μ does not exist, and determination of Π_{CED} values from PVT measurements appears unlikely with the precision required for meaningful χ values.

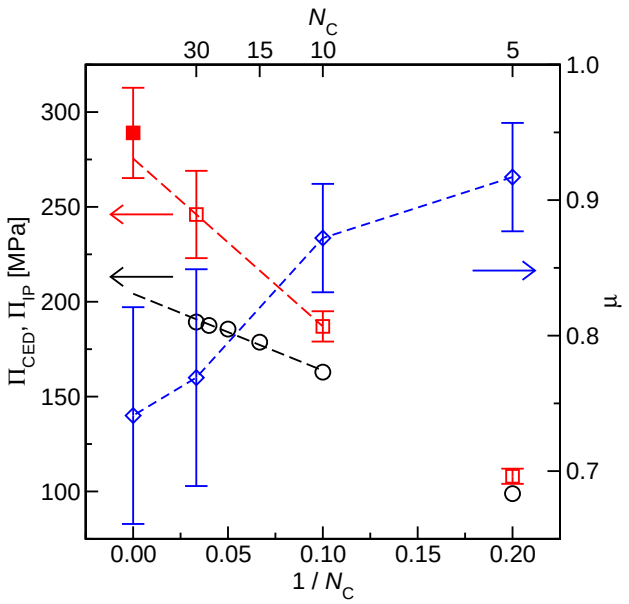

Figure 2.7: Chain-length dependence of Π_{CED}, Π_{IP}, and their ratio, μ, for PEP oligomers at $T = 394$ K. The open black circles, red squares, and blue diamonds represent simulation data for Π_{CED}, Π_{IP}, and μ, respectively. The filled red square shows Π_{IP} from experimental PVT data by Krishnamoorti et al. [26] (red filled square). The blue dashed line is only drawn to guide the eye.

2.3.2 Thermodynamic properties of binary mixtures

Direct Comparison to Experimental Phase Equilibrium Data. To assess the feasibility of the current computational approach, the binary mixture of 2-methylpentane and n-heptane is investigated. For these short chains, experimental data [114] on the mixing thermodynamics can be directly obtained from the vapor–liquid coexistence curve. Figure 2.9 displays pressure–composition phase diagrams and the corresponding composition-dependent separation factors at $T = 318$ and 328 K. The simulation results overestimate the vapor pressure over the entire composition range by around 20%. Note that the simulation data at $T = 328$ K almost coincide with the experimental data [114] at $T = 318$ K. As mentioned previously, the TraPPE–UA force field [95, 96] is

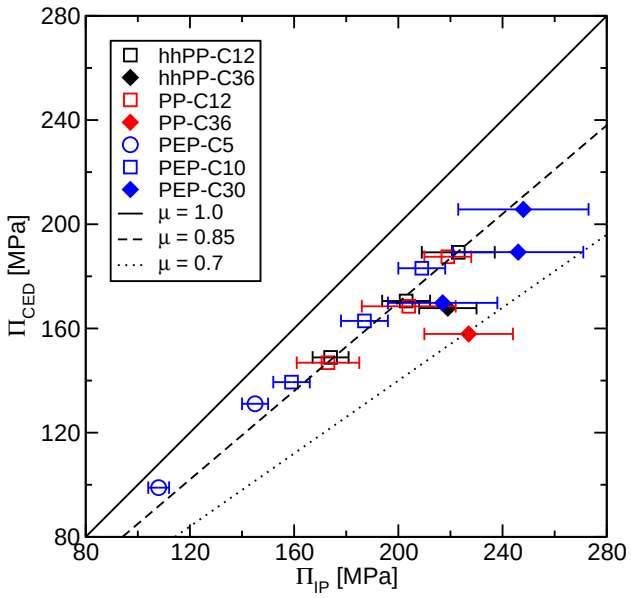

Figure 2.8: Simulated Π_{CED} and Π_{IP} at T = 356, 394 and 440 K for various olefin oligomers at different chain lengths. Note that data at different temperatures are not distinguished by symbol or color for clarity.

known to underestimate normal boiling points of alkanes by about 10 K. However, the separation factors (and relative volatilities) are predicted by the TraPPE–UA force field with high accuracy because the vapor pressures of both compounds are overestimated by a similar amount. Since the free energy of mixing can be computed directly from values on the curve using the assumption that the vapor phase is ideal, the overlap of the two vapor pressure curves indicates that the χ value from simulation at 328 K should be very close to the χ value from experiments at 318 K. Given the weak temperature dependence of χ, this deviation of 10 K is acceptable.

To confirm the agreement of χ values, χ is computed from eqs 2.7 and 2.8, and the results are illustrated in Figure 2.9. In these calculations, the geometric average of the

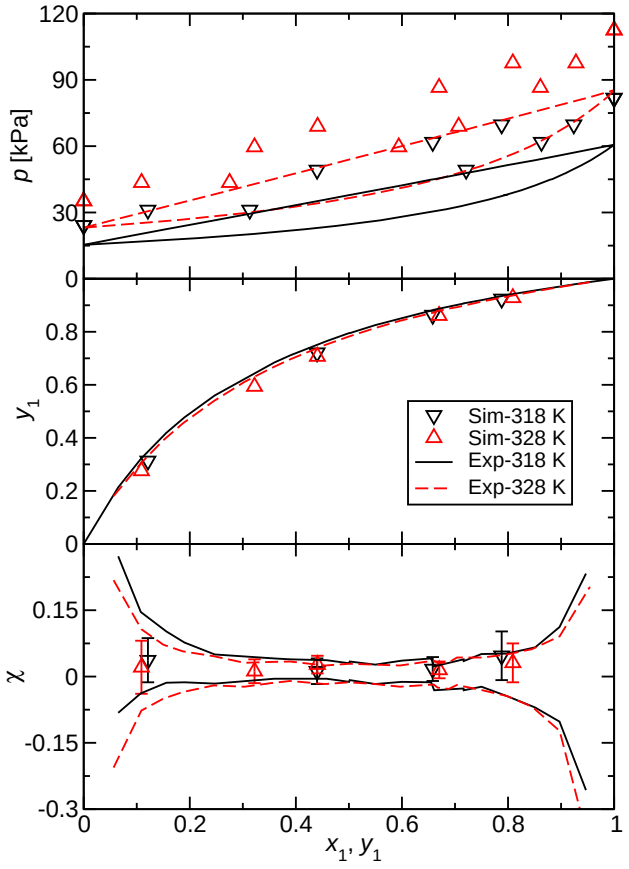

Figure 2.9: Vapor–liquid phase diagram (top), separation factor (middle), and χ parameter as function of composition for the binary mixture of 2-methylpentane (1) and n-heptane (2) at $T = 318$ and 328 K. The experimental data are taken from Berro et al. [114].

molecular volumes of the two components is used as the reference volume. For this binary mixture, the simulations yield a fairly constant, small but positive χ value over the entire composition range. Averaging the simulation data over all compositions yields χ values of 0.028 ± 0.021 and 0.021 ± 0.018 at $T = 318$ and 328 K, respectively, that are in agreement with the experimental values [114] of 0.006 ± 0.012 and 0.017 ± 0.013 at the two temperatures (where the values at the lowest and highest x_1 are not included

in the averages). For both the simulation data and the experimental measurements, the statistical uncertainties in χ increase rapidly as either end of the composition range is approached; this problem is driven by increasing relative errors in gas- and liquid-phase concentrations of the minority species. Overall, the good agreement of predicted and experimentally measured data for this test system confirms that the current computational methodology can be used to study olefin oligomer mixtures.

Oligomeric PP/hhPP Mixture. The vapor–liquid equilibria of equimolar PP and hhPP mixtures are studied at chain lengths $N_C = 12$ and 24. The ΔG_m values are found to closely trace those for an ideal mixture (see Figure 2.10), and the statistical uncertainties for the present simulations are less than 0.02 and 0.1 kJ/mol for the C12 and C24 mixtures, respectively. Figure 2.10 shows a comparison of the χ_{CP} values for these oligomers computed using eq 2.8 to the χ_{SANS} values of PP/hhPP polymer mixtures from SANS experiments [36, 78]. For this system, the χ_{SANS} values are positive and very close to zero, irrespective of temperature and of which of the two monomers is deuterated. The χ_{CP} values calculated from the ratios of vapor densities in the oligomer GEMC simulations are negative, but also quite small and exhibit large statistical uncertainties. The experimental data for the PP/hhPP mixture yield a negative X_E value due to the larger magnitude of the term resulting from the difference in solubility parameters [36, 78]. This stabilized irregular mixing behavior is also observed for the present simulations, but X_E is about twice as large as the experimental value at $T = 394$ and 440 K. Nevertheless, it appears that the simulations for the olefin oligomers yield the qualitatively correct behavior despite their short chain length. Considering also the data at the two higher temperatures, the X_E values appear to decrease slightly in magnitude with increasing temperature, a trend also observed in the experimental data at $T = 356$, 394, and 440 K (with a slope of ≈ 0.001 MPa/K) [36, 78].

Given the large uncertainties in the simulation values and the relatively similar values

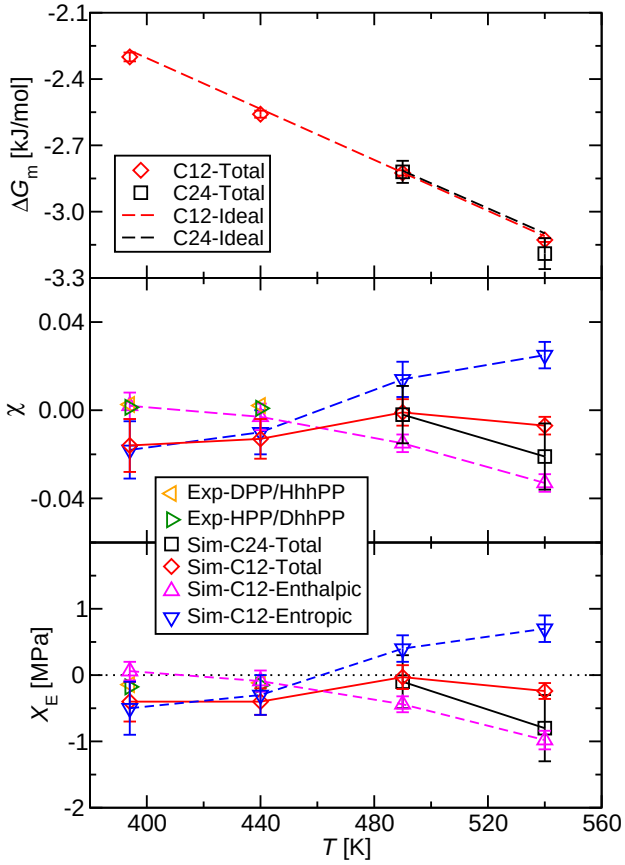

Figure 2.10: Comparison of ΔG_{m} (top), χ (middle), and X_{E} (bottom) values for binary equimolar PP/hhPP oligomer and polymer mixtures at various temperatures. Experimental SANS data [36, 78] are for the polymeric PP/hhPP mixture, where HPP/DhhPP indicates that hhPP was partially deuterated in the mixture, and vice versa. The cyan circles show χ_∞ values estimated from X_{E} and infinite-chain-length solubility parameters from eqs 2.13 and 2.14.

for the C12 and C24 mixtures, it is not possible to ascertain whether deficiencies of the force field or endgroup effects are responsible for the deviation from experiment. Assuming that X_{E} is independent of chain length (see discussion of radial distribution functions), the infinite-chain-length χ_∞ value can be estimated using eq 2.9 and the infinite-chain-length solubility parameters from eqs 2.13 and 2.14. These χ_∞ values are

slightly larger than the oligomer χ values obtained from the GEMC simulations, and χ_∞ values are in better agreement with the experimental data (see Figure 2.10).

To provide additional information on the stabilized irregular mixing behavior, X_E is divided into enthalpic and entropic contributions for the C12 mixture (see Figure 2.10). Despite the negative values for X_E at all temperatures and the weak temperature dependence, the simulations indicate that both the enthalpic and entropic contributions show a significant temperature dependence and may even switch sign between 440 and 490 K. At the higher temperatures (where the statistical errors are smaller), the stabilized mixing appears to result from a larger, favorable enthalpic contribution, whereas the entropic contribution is unfavorable.

To better understand this result, the excess volume, enthalpy, and entropy of mixing are computed for the C12 mixture (see Table 2.2). The PP/hhPP mixture exhibits negative volume change upon mixing, but the average volume change of 0.3 cm^3/mol is only about 0.3% of v_0. Therefore, the FH theory assumption on the incompressibility of the mixture holds well for this system. At the two lower temperatures, $\Delta H_\mathrm{m}^\mathrm{E}$ is close to zero and $\Delta S_\mathrm{m}^\mathrm{E}$ is slightly positive. In contrast, at the two higher temperatures, both $\Delta H_\mathrm{m}^\mathrm{E}$ and $\Delta S_\mathrm{m}^\mathrm{E}$ are negative. $\Delta S_\mathrm{m}^\mathrm{E}$ appears to play a significant role for the PP/hhPP mixture, in contrast to the assumption of FH theory of an ideal entropy of mixing.

Table 2.2: Excess properties upon mixing for the equimolar PP-C12/hhPP-C12 mixture

T [K]	$\Delta V_\mathrm{m}^\mathrm{E}$ [cm^3/mol]	$\Delta H_\mathrm{m}^\mathrm{E}$ [kJ/mol]	$\Delta S_\mathrm{m}^\mathrm{E}$ [J/(K·mol)]
394	-0.22_{10}	0.004_{10}	0.08_6
440	-0.30_8	-0.005_{10}	0.04_4
490	-0.36_6	-0.030_8	-0.06_3
540	-0.57_8	-0.073_{10}	-0.11_2

Structural Analysis. To gain insights on the effects of the liquid structure on the mixing thermodynamics, the carbon–carbon RDFs for the PP-C12/hhPP-C12 equimolar mixture and of pure liquids are analyzed. As can be seen in Figure 2.11, the shapes of the PP–PP and hhPP-hhPP RDFs differ significantly, with the former yielding a higher first than second peak, whereas the reverse is true for the hhPP–hhPP RDF. Interestingly, the RDFs for the binary mixture fall right on top of the corresponding RDFs for pure liquids; previously Heine et al. [49] also reported that mixing does not significantly perturb the molecular structure. The cross PP–hhPP RDF appears to track rather well the average of the PP–PP and hhPP–hhPP RDFs. Since Π_{CED} can be computed from the integration of the RDFs, the similarity between cross-species RDF and the average of two intra-species RDFs implies that the enthalpy of mixing is small. However, the number integrals (obtained by integration of the RDF that gives the number of carbon atoms inside a spherical volume of a certain radius) corresponding to the first peak indicate very small but important changes upon mixing. First, the number integral for the PP–hhPP contacts falls above both those for PP–PP and hhPP–hhPP contacts, instead of being close to their average. Second, the ratio of the hhPP–hhPP number integrals for the equimolar mixture and pure liquid is 0.49, i.e., just slightly below what would be expected from the overall composition. These two features are structural representations of the very slight stabilization upon mixing but also reflect the negative excess volume of mixing.

The temperature and the chain length dependences of the RDFs are illustrated in Figure 2.12. For both PP-C12 and hhPP-C12, the height of the first peak decreases significantly with increasing temperature, and its position shifts to slightly larger distances. For hhPP-C12, the first peak even shrinks into just a shoulder at $T = 540$ K. For PP-C12, the changes in the RDFs with temperature are mostly confined to the first peak, and the RDFs at all four temperatures nearly coincide for $r > 6$ Å. In contrast,

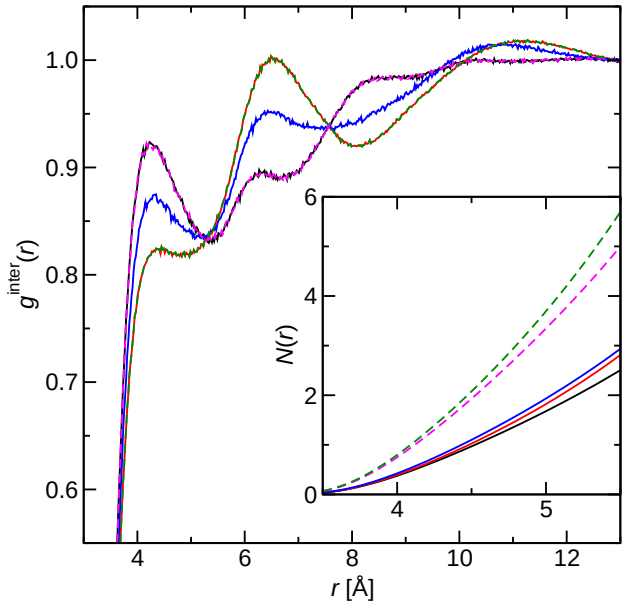

Figure 2.11: Intermolecular carbon–carbon RDFs for C12 oligomers at $T = 440$ K. The dashed magenta and solid black lines denote PP–PP RDFs in the pure melt and the PP-hhPP equimolar mixture, respectively. The dashed green and solid red lines denote hhPP–hhPP RDFs in the pure melt and the PP/hhPP equimolar mixture, respectively. The solid blue line indicates the PP-hhPP RDF in the equimolar mixture. The inset shows the corresponding number integrals for the first peak.

an increase in temperature also leads to a marked decrease in the height of the second peak and the depth of the second minimum for the hhPP-C12 olefin. Both the decrease in structural order and also the decrease in the density (also reflected in the number integrals) lead to a lower Π_{CED} at higher temperatures, as also found for the directly calculated Π_{CED} values at various temperatures.

Increases in the chain length also affect the intermolecular structure (see Figure 2.12), but in a more subtle manner. For PP oligomers, the height of the first peak descreases slightly with increasing degree of polymerization, but its position shifts to slightly shorter distances. In contrast to the temperature dependence, the lowering of

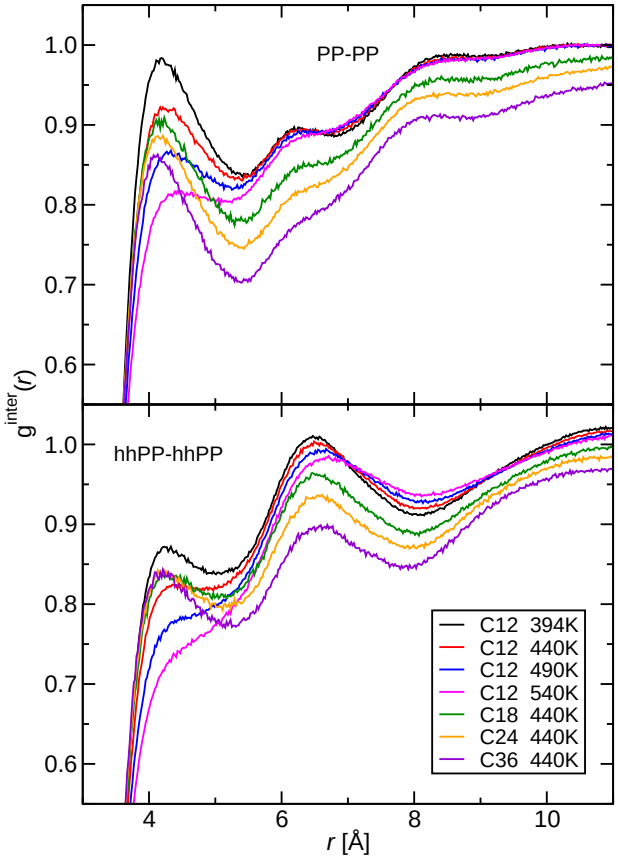

Figure 2.12: Intermolecular carbon–carbon RDFs for pure liquids of PP (top) and hhPP (bottom) oligomers. The black, red, blue, and magenta lines denote RDFs for C12 chains at T = 394, 440, 490, and 540 K, respectively. The green, orange, and purple lines denote RDFs for C18, C24, and C36 chains, respectively.

the intermolecular carbon–carbon RDFs with increasing chain length persists out to about 15 Å for both isomers. For the hhPP oligomers, an increase in chain length does not appear to have a significant effect on the first peak, but leads to a downward shift of the first minimum and a decreased height of the second peak. The downward shift of the RDFs with increasing degree of polymerization indicates some degree of intramolecular backfolding (see below), and accounts for the higher Π_{CED} per monomer observed for shorter oligomers.

The RDFs at different chain lengths can be used to elucidate why the mixing irregularity is not significantly affected by degree of polymerization. Maranas et al. [34] correlated the shapes of pure-component RDFs to the mixing behavior. In their work, d is defined as the height difference between the first maximum and the first minimum of the single-component RDF, and Δd is defined as the difference in d for a binary pair. Maranas et al. [34] found that the magnitude of X_{E} increases with the magnitude of Δd. Based on this correlation, d and Δd for PP and hhPP of different chain lengths are computed at $T = 440$ K, and results are summarized in Table 2.3. Despite the chain-length dependence of the RDF shapes, Δd remains constant for all chain lengths in this study, thereby indicating that already relatively short oligomers may capture the structural subtleties responsible for irregular mixing behavior.

Table 2.3: Structural characteristics d and Δd for PP and hhPP oligomers $T = 440$ K.

N_{C}	d_{PP}	d_{hhPP}	Δd
12	0.09	0.01	0.08
18	0.12	0.03	0.09
24	0.14	0.05	0.09
36	0.15	0.07	0.08

With respect to chain conformation, the squared end-to-end distance, R^2, and the statistical segment length, b, are important characteristics. The latter quantity is calculated via $b = \sqrt{R^2/N}$, where N is the degree of polymerization, defined here as $N = N_{\mathrm{C}}/6$. The calculated R^2 values are found to increase approximately as $N_{\mathrm{C}}^{3/2}$ because the oligomers considered here are too short to conform to the Flory scaling of the chain dimension with degree of polymerization expected for polymers in a melt ($\nu = 1/2$). However, the b values show clear convergence toward the infinite-chain-length limit. The b values for the PP-C36 and hhPP-C36 olefins at $T = 440$ K are 7.00 ± 0.06 and 7.41 ± 0.04 Å, respectively. Assuming that b converges to the infinite-chain-length

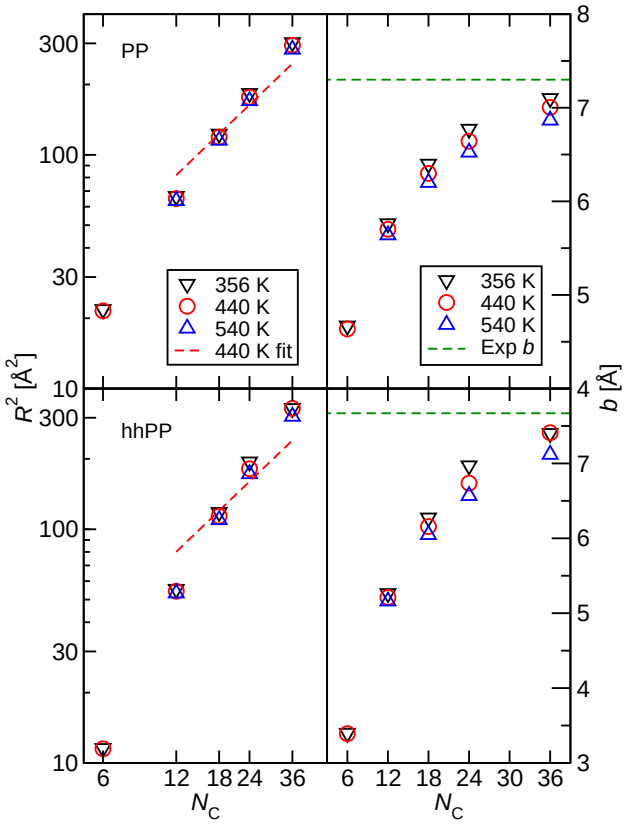

Figure 2.13: The average squared end-to-end distance, R^2 (left column), and the statistical segment length, b (right column), for PP (top row) and hhPP (bottom row) oligomers in their pure liquids at various temperatures (black down triangles, red circles, and blue up triangles for $T = 356$, 440, and 540 K, respectively). The red dashed line indicates a fit using $\nu = 1/2$ for $12 \leq N_C \leq 36$. The green dashed lines indicate the experimental b values for polymers obtained by Graessley et al. [36], that were found to be indistinguishable over the temperature range from 300 to 440 K.

value as function of $N^{-3/2}$ (and using only data for C12 to C36 chains) yields b_∞ values of 7.2 ± 0.1 and 7.8 ± 0.2 for PP and hhPP polymers that agree well with the corresponding experimental values [36] of 7.30 ± 0.08 and 7.67 ± 0.05, respectively. Again it is evident that these oligomers already reflect most of structural details of their polymeric counterparts. The decrease in R^2 and b values with increasing temperature indicates higher

flexibility and a stronger tendency to form coils (or partially fold back for these short oligomers). These changes in chain conformation reduce the number of intermolecular contacts and contribute to the reduction of Π_{CED} at higher temperatures.

2.4 Conclusions

In this work, MC simulations are carried out to investigate the liquid phase and vapor–liquid equilibria for olefin oligomers and their mixtures. The compounds are selected to represent low-molecular-weight analogues of polypropylene, head-to-head polypropylene, and poly(ethylene-*alt*-propylene). Simulations using the united-atom version of the TraPPE force field for the oligomers, where unlike Lennard-Jones interaction parameters are determined by the Lorentz-Berthelot combining rule, and extrapolation to the infinite-chain-length limit yields excellent agreement for the liquid densities of pure melts and for the differences in solubility parameters, δ, between pairs of polymers where δ values are estimated directly via the internal energies for the simulation or from SANS experiments. The deviation from oligomer experimental data can be attributed more to the choice of force field than the simulation methodology. With further increases in algorithmic efficiency and computer power, simulations using explicit-hydrogen models and/or non-pairwise additive interactions may become achievable for these oligomer systems. The simulations also confirm that cohesive energy densities obtained via the internal pressure route significantly overestimate the correct values determined via the internal energy route. The degree of overestimation is found to depend on chain length, chain architecture, and temperature. This implies that experimental solubility parameter data for polyolefins that are approximated from Π_{IP} should be used with caution, and that force-field-based simulations offer a more accurate (and also precise) approach to obtain Π_{CED} values and the corresponding solubility parameters of polyolefins. However, the persistence length needs to be considered when deciding on the lower and upper

bounds of the oligomer length range included for the extrapolation of Π_{CED} values to the long-chain limit.

The free energies of mixing for binary mixtures of olefin oligomers can be calculated directly from the Gibbs free energies of transfer, i.e., ratios of number densities. A model binary system of short chains proves the feasibility of this method and yields FH χ parameters that match values calculated from experimental phase equilibrium data. However, the statistical uncertainties for both simulation and experiment are quite large. When this approach is applied to PP/hhPP mixtures with 12 or 24 carbon atoms, then the simulations are able to reproduce (albeit with large uncertainties) the stabilized irregular mixing behavior observed in SANS experiments for the corresponding polymers. At higher temperatures, the mixing irregularity is due to a favorable enthalpic contribution, but there is some indication that entropic contributions stabilize mixing at lower temperatures.

Structural analysis indicates that mixing does not strongly affect the like–like radial distribution functions, but there is a small increase in unlike contacts compared to an ideal mixture. Some structural features, particularly, the height difference between the first peak and the first minimum in the RDFs are well preserved for all chain lengths and may explain why the irregular mixing behavior is not strongly dependent on the degree of polymerization.

Overall, the study in this chapter indicates that simulations of oligomer mixture have much to offer for predicting and understanding the mixing behavior of polymers. Future work will need to address whether this methodology can also be applied to less ideal blends involving poly(*iso*-butylene) and repeat units with aromatic rings or more polar functionalities.

Chapter 3

Understanding the Molecular Weight Dependence of χ and the Effect of Dispersity on Polymer Blend Phase Diagrams

3.1 Introduction

The ability to engineer diverse polymers on the molecular level enables the creation of materials with a wide range of properties. Polymer blends, a soft matter analogue of alloys, provide a means to access materials properties that are beyond the reach of single-component polymeric materials [2, 11, 115, 116, 117]. Knowledge of how the interactions between unlike segments influence the phase behavior is the key to the custom tailoring

This work was in collaboration with Shuyi Xie and Reza Foudazi, and this chapter is reproduced in part with permission from (Q. P. Chen, S. Xie, R. Foudazi, T. P. Lodge and J. I. Siepmann, "Understanding the Molecular Weight Dependence of χ and the Effect of Dispersity on Polymer Blend Phase Diagrams", *Macromolecules* **2018**, DOI: 10.1021/acs.macromol.8b00604, American Chemical Society)

of desired polymeric materials properties. Flory–Huggins (FH) theory [24, 25] is widely used to correlate the mixture phase behavior with a binary interaction parameter χ:

$$\frac{\Delta G_{\mathrm{m}}}{k_{\mathrm{B}}T} = \frac{\phi_1}{N_1}\ln(\phi_1) + \frac{1-\phi_1}{N_2}\ln(1-\phi_1) + \chi(N_1, N_2, T)\phi_1(1-\phi_1) \tag{3.1}$$

where ΔG_{m} is the free energy of mixing per reference site; k_{B} and T are the Boltzmann constant and the absolute temperature, respectively; N_i (i = 1, 2) is the ratio between molar volume of component i and the reference volume; and ϕ_1 is the volume fraction of component 1. Experimentally, χ is typically fitted to either the binodal curve [118, 119] using eq 3.1 (denoted as χ_{eff}), or to the structure factor from neutron scattering experiments [75] using the random-phase approximation [48] (denoted as χ_{SANS}).

From eq 3.1, accurate determination of χ is central to predicting the polymer blend phase diagram. In addition to theoretical advances [39, 41, 43], breakthroughs in computational method development enable a drive towards predicting χ from the molecular structures of polymers. One method is to perform molecular simulations on miscible binary blends, compute the structure factor, and then fit χ using the random-phase approximation, which mimics how χ_{SANS} is determined from neutron experiments [49, 50]. Another approach is to obtain χ through interfacial concentration profiles from immiscible binary blend simulations [120, 71]. In contrast to the above two methods where substantial assumptions are introduced in the calculation, vapor–liquid equilibria simulations are utilized to compute χ from the chemical potential in my earlier work [121] (denoted as χ_{CP}). Recently, Zhang et al. developed a new method to compute χ by performing thermodynamic integration on the path along which one polymer transforms into the other [63].

Despite this progress, two major questions still remain on how the coupling of eq 3.1 and the χ calculation method can be used to achieve the prediction of phase diagrams

for binary polymeric mixtures. The first problem is regarding the molecular weight dependence of χ. Since χ is often calculated for one specific set of molecular-weights, it is important to know its molecular weight dependence in order to extrapolate χ to the molecular weights of interests. Unfortunately, the exact functional form of $\chi(N_1, N_2)$ for polymer blends is still a matter of discussion. For example, Han et al. studied polystyrene/poly(vinyl methyl ether) blends and concluded that χ_{SANS} for three sets of molecular weights that vary by about a factor of three are indistinguishable [28]. In contrast, Nedoma et al. reported that for polyisobutylene/(deuterated polybutadiene) blends, $\chi_{\mathrm{SANS}} \sim 1/N_{\mathrm{AVE}}$, where $N_{\mathrm{AVE}} = 4(N_1^{-1/2} + N_2^{-1/2})^{-2}$ [31, 122]. A similar dependence of $\chi(N) \sim N^{-1}$ is also observed for symmetric diblock copolymers [123]. A theoretical study by Morse and Chung predicted a $\chi \sim N^{-1/2}$ dependence for polymer "mixtures" with $\chi \approx 0$. They attributed this relationship to the enhanced intra-chain interactions as the chain length increases and thus the screening of the inter-chain interactions that contribute to χ [124]. In addition to the findings above, it is also intriguing whether the molecular weight dependence of χ implies the failure of the FH theory, since χ is a molecular-level parameter and should be independent of chain length [125].

The second problem is whether eq 3.1 is sufficiently accurate for a disperse "binary" mixture (i.e., a quasi-binary mixture), where molecular weight distribution (MWD) and dispersity (Đ) might also play a role. Experimental phase diagrams of disperse polystyrene in methylcyclohexane show that coexistence curves for different total compositions do not collapse on one another, and the coexistence curves of disperse samples deviate by a few Kelvin from those of the relatively monodisperse ones [126, 127]. For a bimodal polymer distribution when the chain length ratio between the two polymer species is sufficiently large, the solution can phase-separate into three coexisting liquid phases near the critical point [128, 129, 130]. Such deviations from binary mixture phase behavior motivate theoretical developments to understand the effect of dispersity. The

standard approach is to apply continuous thermodynamics to the FH theory to calculate the cloud point curve of the quasi-binary mixture [131, 132, 133, 134, 135, 136, 137, 138]. In addition, the partitioning of different molecular weight species between coexisting phases is identified by both experimental [139, 140] and theoretical [134, 138, 141, 142] studies.

The key barrier in studying these two problems is the limited molecular weight range, MWD, and Đ accessible via experiments. On the one hand, a wide range of molecular weights is essential to provide sufficient precision in the determination of $\chi(N)$. On the other hand, however, the molecular weight range is restricted by the narrow temperature window for scattering or phase diagram measurements, which is limited between the glass transition temperature and the degradation temperature of polymers. Similarly, for the effects of MWD and Đ, it is challenging to prepare polymer samples with the desired MWD and Đ to achieve statistically meaningful conclusions.

Molecular simulations provide a convenient route to overcome the two problems mentioned above. First, simulations can access a much broader range of temperature without degradation. Second, both MWD and Đ can be tuned with relative ease in the simulations. Third, simulations can directly access compositional and structural information that is difficult to extract from experiments, which can shed light on the molecular origin of the phase behavior.

Despite these advantages, direct simulations of polymer mixture phase behavior are rare, primarily due to the sampling difficulty of molecular transfers between two phases (needed for the computation of χ as well as the phase diagram). For example, the effect of dispersity on phase equilibria has only been simulated for Lennard-Jones particles [143, 144, 145]. Fortunately, a variety of sampling techniques have been developed to assist the transfer of large molecules in order to realize the routine application of polymer (oligomer) phase equilibria simulations [53, 54, 55, 56, 64, 65, 57, 146, 58, 59, 60, 61, 62]

In addition to the sampling difficulty, another challenge is the high sensitivity of the phase diagram to the free energy of transfer that governs the phase separation, which poses a stringent test on the accuracy of the molecular model used in the simulations. As shown in a recent example of the alkane/alkanol mixture phase diagram [62], a difference of 1 kJ/mol in the transfer free energy, often regarded as a "small" deviation in force field development, results in about 50 K deviation in the coexistence curve.

In this work, Gibbs ensemble Monte Carlo (GEMC) simulations [92, 93] with advanced sampling techniques in conjunction with experimental cloud point measurements are utilized to study the phase behavior of model mixtures consisting of oligomeric poly(ethylene-*alt*-propylene) (PEP) and poly(ethylene oxide) dimethyl ether (PEO). Figure 3.1 shows the repeat unit structures of the two molecules. This incompatible mixture is selected as the model system because of the relatively low molar mass ($n < 10$) needed to demix the blend at room temperature, which permits the determination of phase diagrams by both simulations and cloud point measurements. In this chapter, results from simulations and experiments regarding χ of PEP/PEO mixtures with various molecular weights are presented, from which the molecular weight dependence of χ is identified and rationalized. In the second part, coexistence curves for quasi-binary mixtures consisting of monodisperse PEP and disperse PEO are first presented. The MWD and Đ of PEO are varied but the same M_{n} of 500 Da is maintained to understand the effects of MWD and Đ on the phase diagram. Further analysis on the partitioning of different molecular weight PEO molecules and structures of the liquid in the two coexisting phases provides additional insight into the phase behavior of this mixture.

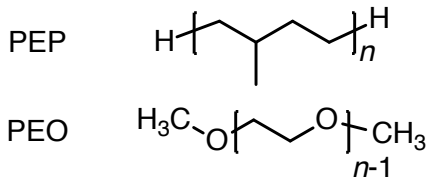

Figure 3.1: Repeat unit structures of PEP and PEO oligomers studied in this work.

3.2 Methods

3.2.1 Nomenclature.

A range of PEP and PEO chain lengths are studied to explore the molecular weight dependence of χ, and they are denoted as [molecule]-[M_n]-[MWD]-[Đ], in which M_n is the number average molecular weight. MWD and Đ in the nomenclature may be omitted if the sample is monodisperse. The MWD studied includes "MD" (monodisperse), "BM" (bimodal), and "SZ" (Schulz-Zimm) [147]. For example, PEO-222 indicates monodisperse oligomer PEO with five repeat units ($n = 5$), while PEO-500-SZ-1.10 represents PEO with an M_n of 500 Da, a Schulz-Zimm MWD, and a dispersity of 1.10. The number of repeat units n, as defined in Figure 3.1, is also used to indicate the chain length of PEP and PEO. The molecular weight information of PEO used in the simulations can be found in Figure 3.2. Note that PEP*-423 in this work refers to the squalane (2,6,10,15,19,23-hexamethyltetracosane, similar to the PEP-423 except for a head-to-head connectivity between the two middle segments) for the convenience of comparison with experimental data.

3.2.2 Simulation details.

All the simulations were performed using the in-house Monte Carlo (MC) simulation software package MCCCS–MN (Monte Carlo for Complex Chemical Systems–Minnesota)

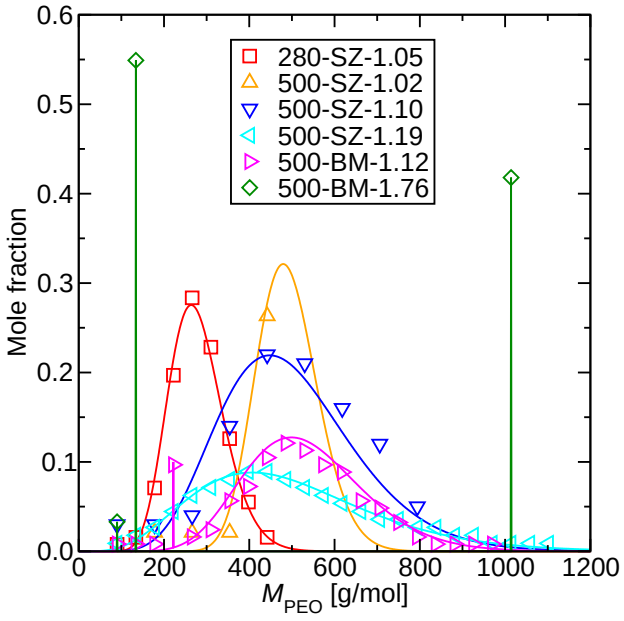

Figure 3.2: Molecular weight distribution of disperse PEO samples used in the simulations. Symbols represent simulation data, and lines are the Schulz-Zimm or bimodal distribution fit to the symbols.

[99]. GEMC simulations [92, 93] in the canonical (NVT) or isobaric-isothermal (NpT) ensemble were used to simulate the vapor–liquid equilibira (VLE) of single-component oligomers and their binary mixtures. For single-component oligomer simulations, 300 molecules were used for PEO-266 ($n = 6$), and 400 molecules were used for other oligomers. For binary mixtures, the PEP weight fraction was around 50%, and the system sizes were varied to allow for at least 10% of each molecule type in the vapor phase and a total system size of no smaller than 3000 interaction sites. The detailed system sizes for each system are listed in Table 3.1.

A similar simulation protocol was used as described in Chen et al. [121]. In brief, all the molecules were placed in the liquid phase during the initialization of the system. Center-of-mass translations, center-of-mass rotations, conformational moves [96],

Table 3.1: Overall number of molecules for each component in VLE simulations.

System	T [K]	PEP-M			PEO-M				
		72	142	212	90	134	178	222	266
PEP-142	435/490	2	400	0	0	0	0	0	0
PEP-212	435/490	2	2	400	0	0	0	0	0
PEO-134	435/490	0	0	0	2	400	0	0	0
PEO-178	435/490	0	0	0	2	0	400	0	0
PEO-222	435/490	0	0	0	2	0	2	400	0
PEO-266	435/490	0	0	0	2	0	2	0	300
PEP-142/PEO-134	435/490	0	150	0	0	150	0	0	0
PEP-212/PEO-134	435/490	0	0	75	0	250	0	0	0
PEP-142/PEO-178	435/490	0	180	0	0	0	120	0	0
PEP-212/PEO-178	435/490	0	0	100	0	0	150	0	0
PEP-142/PEO-222	435	0	700	0	0	0	0	100	0
PEP-142/PEO-222	490	0	220	0	0	0	0	100	0
PEP-212/PEO-222	435/490	2	2	200	2	0	2	200	0
PEP-212/PEO-266	435	2	2	200	2	0	2	0	85
PEP-212/PEO-266	490	2	2	150	2	0	2	0	85

volume exchange moves, and particle transfer moves were used to sample the configurational phase space of the system. In order to aid the molecular transfer between the vapor and the liquid phases, impurity molecules of shorter PEP or PEO oligomers (two molecules each) were also used, together with the inter-box identity switch moves [100, 102]. This approach boosts the acceptance of transfer moves [102, 105] without sacrificing the accuracy, if proper corrections are applied ($Đ < 1.01$) [121].

All VLE simulations were equilibrated for around 1.5×10^5 MC cycles (MCCs, consisting of N randomly selected moves, where N is the total number of molecules) until there was no drift in energies and number densities of each molecule in both phases. Then, simulations were run for at least another 3×10^5 MCCs as the production stage. The k-d tree data structure was used to accelerate unary PEP simulations [148]. Single-component thermodynamic properties such as liquid densities (ρ), vapor pressures (p_{vap}),

cohesive energy densities (Π_{CED}), and solubility parameters (δ), as well as the χ parameters for binary mixtures, were extracted from these simulations as described in the previous chapter [121]. In particular, Π_{CED} and δ were calculated using the following equation:

$$\delta = \sqrt{\Pi_{\mathrm{CED}}} = \sqrt{\frac{U_{\mathrm{vap}} - U_{\mathrm{liq}}}{V_{\mathrm{liq}}}} \tag{3.2}$$

where U_{vap} is the molar internal energy of the molecule in the vapor phase, U_{liq} is the molar internal energy of the liquid phase, and V_{liq} is the molar volume of the liquid phase.

In addition, liquid–liquid equilibria (LLE) of PEP*-423/PEO mixtures were simulated via NpT-GEMC simulations ($p = 1$ bar). A three-box setup was used, in which two simulation boxes were used to represent two coexisting liquid phases while the third simulation box was used as the transfer medium [149]. A total of about 120 PEP molecules and more than 100 PEO molecules of various molecular weights were used in the simulations, which led to 40–60% overall PEP weight fraction. The same types of moves as used in the VLE simulations were also applied to the LLE simulations. In addition, self-adapting fixed-end-point configurational-bias Monte Carlo (SAFE-CBMC) moves [106] were applied to PEO molecules when $n_{\mathrm{PEO}} \geq 10$, to ensure good sampling of the internal conformation of these long-chain molecules. Furthermore, shorter oligomeric PEP molecules were used as impurities (e.g., $n_{\mathrm{PEP}} = 1$–5). This resulted in a dispersity of 1.04 for PEP*-423. Impurities were also used for PEO but they were included in the MWD and the calculation of Đ. Similar to VLE simulations, direct particle transfer moves were only used for the shortest oligomers (i.e., PEP-72 and PEO-90), and the transfer of higher molecular weight molecules was achieved via the inter-box identity switch moves, which were applied to all the neighboring molecular weight oligomer pairs (e.g., PEP-72 and PEP-142). The free energy of molecular transfer to and from the vapor phase was biased to ensure that there are on average 0.5–1.5 molecules for

each PEP impurity molecule and 0.2–1.5 molecules for each PEO molecule in the vapor phase. Such bias does not affect the resulting liquid–liquid phase diagram because the same biasing potential is applied to both liquid–vapor box pairs and its net effect on the liquid–liquid free energy of transfer is thus zero [62]. PEO molecules with $n \geq 19$ were not allowed to transfer to the PEP-rich liquid box or the vapor box because of their negligible solubilities in the alkane phase even at the highest temperature studied. Similar to VLE simulations, all the systems were initialized with all the PEP molecules in one liquid box and all the PEO molecules in the other liquid box. These LLE simulations were equilibrated for at least 2×10^6 MCCs until there was no drift in energies or compositions for each phase, and the production lasted for at least another 5×10^5 MCCs.

Eight and 16 independent simulations were performed for VLE and LLE simulations, respectively. Statistical uncertainties of the simulation data were estimated from these uncorrelated runs and are reported as the 95% confidence interval.

3.2.3 Molecular models.

The TraPPE–UA (transferable potentials for phase equilibria–united atom) force field [95, 96, 150] was used to model alkanes and ethers investigated in this study. The force field treats these molecules as pseudo-atoms connected by the prescribed bond length, and bending angles and dihedral angles are allowed to vary. One CH_x (x = 1, 2, and 3) group or one oxygen atom is described as one pseudo-atom. The Lennard-Jones (LJ) 12-6 potential with the Lorentz–Berthelot combining rule [97] and the Coulomb potential for partial charges on the oxygen and α-carbon sites were used to model the interactions between two pseudo-atoms. A spherical cutoff r_{cut} of 14 Å was used, and the interactions beyond the truncation distance were accounted for via analytical tail corrections [98]. The Ewald summation method [149] with a screening parameter of

$\kappa = 3.2/r_{\text{cut}}$ and $K_{\text{max}} = \text{int}(\kappa L_{\text{box}}) + 1$ (L_i is the box length) was used to calculate the Coulomb interactions in the simulations involving PEO molecules.

In addition, the TraPPE–UA ether force field was modified to model the interactions between PEP and PEO more accurately. The original TraPPE–UA force field [150] can reproduce the single-component properties of PEO oligomers, as demonstrated in Figure 3.3, where the vapor–liquid equilibria data were computed for an ether dimer, 1,2-dimethoxyethane, or PEO-90 from our nomenclature, but it fails to accurately predict χ for the binary alkane/ether mixture. As shown in Figure 3.4, the χ parameter of the *n*-dodecane/PEO-178 mixture is overestimated by more than 75% over the entire composition range for the original TraPPE–UA force field (magenta), which implies that the binary interaction between alkane and ether is too unfavorable. The overestimation can be remedied by reducing the partial charges on the PEO atoms and adjusting the LJ parameters of the oxygen atom (i.e., rebalancing the dispersive and Coulomb contributions to the cohesive energy of PEO molecules). Since ether molecules can interact with alkane molecules through dispersive interactions but not dipole interactions in this non-polarizable model, this approach can result in more favorable alkane-ether interactions, and equivalently, a smaller χ. Similarly, other literature findings also suggest that reducing the partial charges on the TraPPE–UA ether molecules enables more accurate prediction of the ether-water phase diagram [61] and Kovats retention indices for alkanols in PEO-type stationary phases [155].

Therefore, I modified the TraPPE–UA ether force field by scaling down the partial charges on the ether molecules and adjusting the LJ parameters of the oxygen atom. Several values of the oxygen partial charge were attempted between the original -0.50 e to -0.44 e. This range was selected because an earlier study suggested that a charge value around -0.45 e is optimal for ether-water interactions [61]. The partial charge on the oxygen atom determines the partial charge on the α-carbon atoms to maintain the

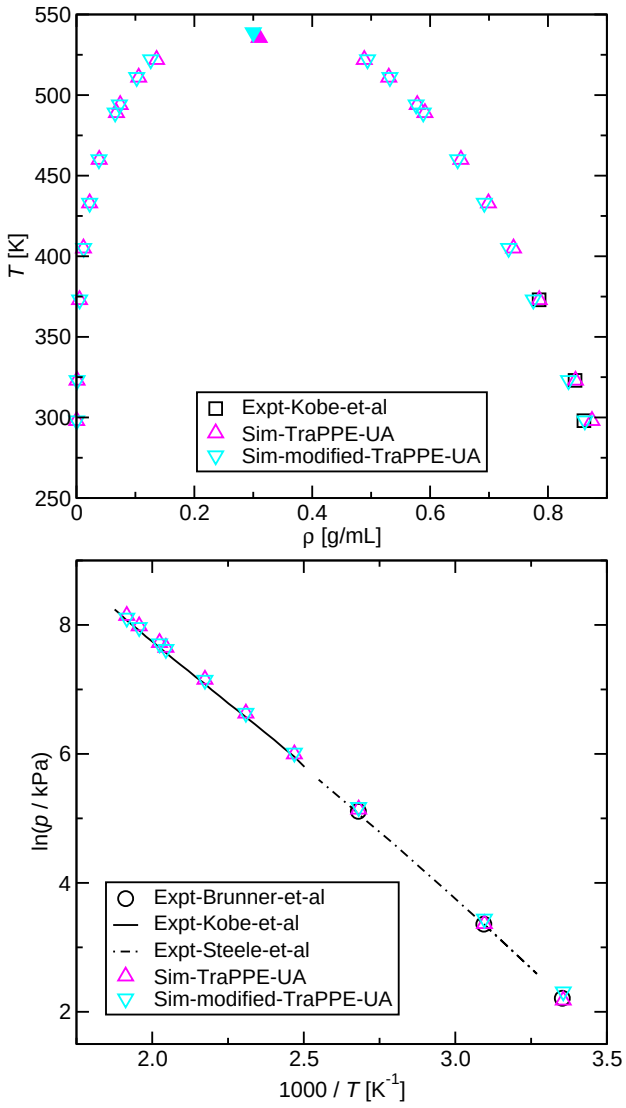

Figure 3.3: (a) Vapor–liquid coexistence curve for 1,2-dimethoxyethane (PEO-90); (b) Saturated vapor pressure as a function of inverse temperature. Experimental data are from Kobe et al. [151], Steele et al. [152], and Brunner et al. [153]. The experimental critical temperature is 539.2 ± 0.4 K [154].

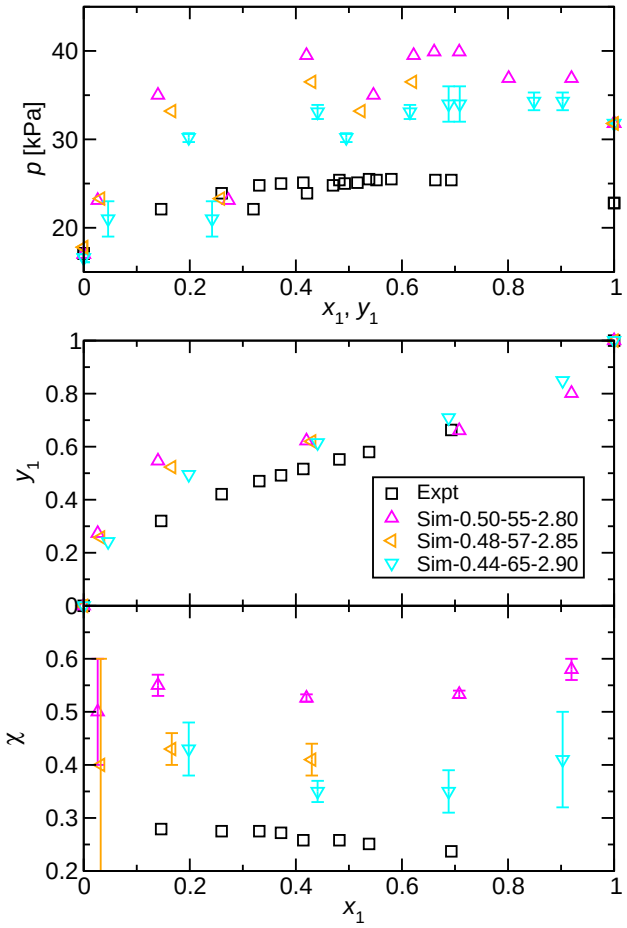

Figure 3.4: Vapor–liquid phase diagram (top), vapor composition as a function of liquid composition, *i.e.*, separation factor (middle), and χ parameter (0.1 nm^3 is used as the reference volume) as a function of liquid composition for the binary mixture of *n*-dodecane (denoted as "1") and triethylene glycol dimethyl ether (PEO-178 from our nomenclature). The legend indicates q_O, ϵ_O/k_B, and σ_O (in units of e, K, and Å), respectively. Experimental data are taken from Treszczanowicz and Cieślak [156].

charge neutrality of the molecule. In addition, the LJ parameters of the oxygen atom were also tuned to reproduce the single-component thermodynamic properties such as liquid densities, vapor pressures, and the critical point of a representative ether molecule, namely PEO-90. For each partial charge value, χ_{CP} for *n*-dodecane/PEO-178 mixtures was computed to allow for the selection of an optimal partial charge value.

It was found that χ_{CP} for the *n*-dodecane/PEO-178 mixture decreases with decreasing partial charge on the oxygen atom, and $q_O = -0.44\ e$ worked best for this system among the partial charge values we selected (see Figure 3.4). The LJ parameters used for this partial charge are $\epsilon/k_B = 65$ K and $\sigma = 2.90$ Å (shown in Table 3.2). Note that much smaller statistical uncertainties for the original ether force field data set (in magenta) are due to a much longer simulation trajectory. Indeed, one can see from Figure 3.5 that χ_{CP} is reduced by a factor of 1.5, from around 0.526 ± 0.007 to 0.35 ± 0.02 (when $x_1 \approx 0.42$–0.44, $v_{ref} = 0.1$ nm^3). The deviation from the experimental χ_{CP} is reduced from more than 100% to around 35%, while the single-component properties of PEO-90, such as liquid densities, vapor pressures, and the critical point are still accurately reproduced. From Figure 3.3, the mean unsigned percentage error (MUPE) for three low-temperature liquid densities is 0.9% (original TraPPE–UA gives 0.6%), for unary vapor pressures is 7% (original TraPPE–UA gives 6%), and for the critical temperature is within 0.1% (original TraPPE–UA gives 0.7%).

Table 3.2: Nonbonded LJ parameters and partial charges of the modified PEO force field.

Atom type	ϵ/k_B [K]	σ [Å]	q [e]
CH_3	98	3.75	0.22
CH_2	46	3.95	0.22
O	65	2.90	−0.44

Moreover, to show that the modification of alkane force field does not impact χ, one

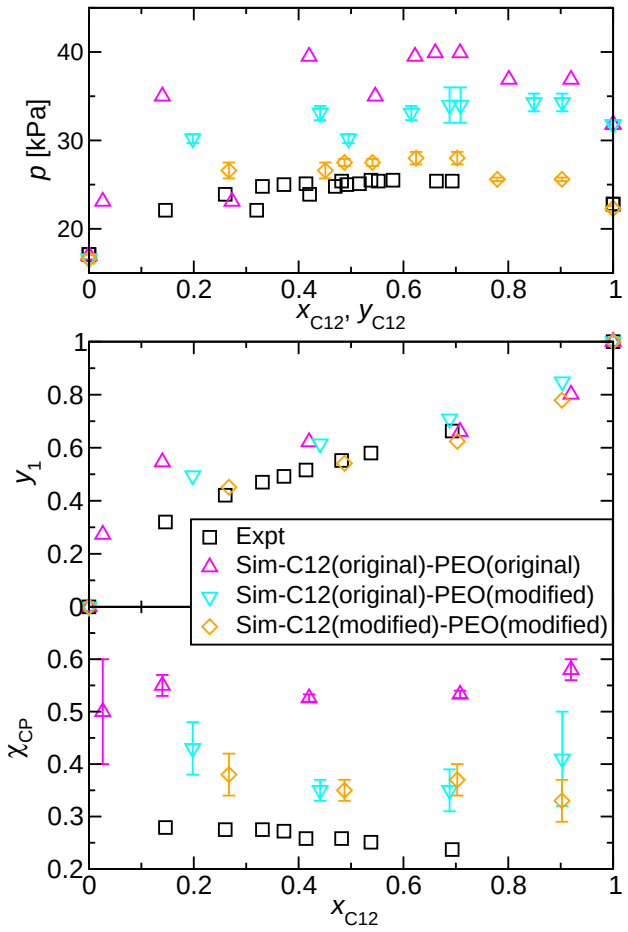

Figure 3.5: Vapor–liquid phase diagram (top), vapor composition as a function of liquid composition, i.e., separation factor (middle), and χ parameter (0.1 nm^3 is used as the reference volume) as a function of liquid composition for the binary mixture of *n*-dodecane (denoted as "C12") and triethylene glycol dimethyl ether (denoted as "PEO"). The data are from different force field combinations, and we refer readers to the text for more details. Experimental data are taken from Treszczanowicz and Cieślak [156].

set of simulations was performed with scaled ϵ values for all the interaction sites in the alkane molecules ($\epsilon_{CH3}/k_B = 100.9$ K and $\epsilon_{CH2}/k_B = 47.4$ K). This set of ϵ reduces unary vapor pressures of n-dodecane to match the experimental ones, but note that this fix can lead to significant deviations in liquid densities and the critical temperature of unary alkanes. One can see in Figure 3.5 that despite much better agreement on the separation factor for the binary system (from cyan to orange in the top and middle panels), χ_{CP} remains approximately the same and shows little improvement. This finding suggests that χ only reflects the binary interactions between two unlike monomers regardless of the deviation in single-component vapor pressures. Since this crude modification of the alkane force field can result in inaccurate predictions of important thermodynamic properties such as liquid densities and the critical temperature, the modified TraPPE–UA ether force field was used to model PEO for the rest of this study (see Table 3.2) while the original TraPPE–UA force field was used to model alkanes.

3.3 Results and Discussion

3.3.1 The molecular weight dependence of χ

Phase diagram of PEP*-423/PEO mixtures. Figure 3.6 shows simulated coexistence curves from LLE simulations and experimental cloud point curves (by courtesy of Shuyi Xie and Reza Foudazi) for quasi-binary blends consisting of PEP*-423 and PEO with various M_n values ranging from 217 to 500 Da. Due to the relatively low dispersities of the PEO samples, the binary FH theory with $\chi(T)$ as the only adjustable parameter (eq 3.1) can be utilized to fit the binodal curves and extract $\chi_{eff}(T)$ [157, 131]. Solid lines in Figure 3.6 are the fitted binodal curves obtained by using a molecular weight-dependent but composition-independent functional form of $\chi_{eff}(T) = a(M)/T + b(M)$.

The fitting results agree fairly well for both simulation coexistence curves and experimental cloud point curves. This suggests that the compositional dependence of χ can be neglected for PEP/PEO mixtures, which is consistent with earlier experimental studies on this system [118, 119] as well as the n-dodecane/PEO-178 simulation data shown in the previous section.

A comparison between simulation and experimental phase diagrams reveals that simulations overestimate the upper critical solution temperatures (UCSTs) by around 50 K. The overestimation in χ is the origin of this deviation. For example, the χ_{eff} values for $M_{\text{n}} = 500$ Da at $T = 450$ K are 0.30 and 0.39 for experimental and simulation data, respectively ($v_{\text{ref}} = 0.1$ nm^3). This 35% overestimation (~ 0.1 in absolute magnitude) is in agreement with the 35% deviation (0.1 in absolute magnitude) in χ_{CP} for the n-dodecane/PEO-178 mixture. Despite the seemingly significant deviation, a 50 K overestimation in UCST corresponds to less than 1 kJ/mol in the free energy of transfer [62]. As a comparison, the original TraPPE–UA force field likely overestimates the UCST by more than 100 K (at $T = 550$ K, the squalane weight fractions in the coexisting phases for the original force field are 88 ± 2 and 4 ± 1 wt%, while this temperature is already above the UCST of the modified force field used here). Although the PEP/PEO interactions are still slightly too unfavorable even after the force field modification, the discrepancy between simulation and experimental results is sufficiently small that reliable results can be obtained regarding the effect of molecular weight and MWD on χ.

For both simulation and experimental data, the fitted χ_{eff} increases with increasing PEO molecular weight. Furthermore, the theoretical phase diagrams for PEP*-423/PEO-222(PEO-217) and PEP*-423/PEO-280 mixtures were calculated using the $\chi_{\text{eff}}(T)$ deduced from PEP*-423/PEO-500-SZ-1.1 and are shown as dashed lines in Figure 3.6. One can see that if χ_{eff} is assumed to be molecular weight independent, the

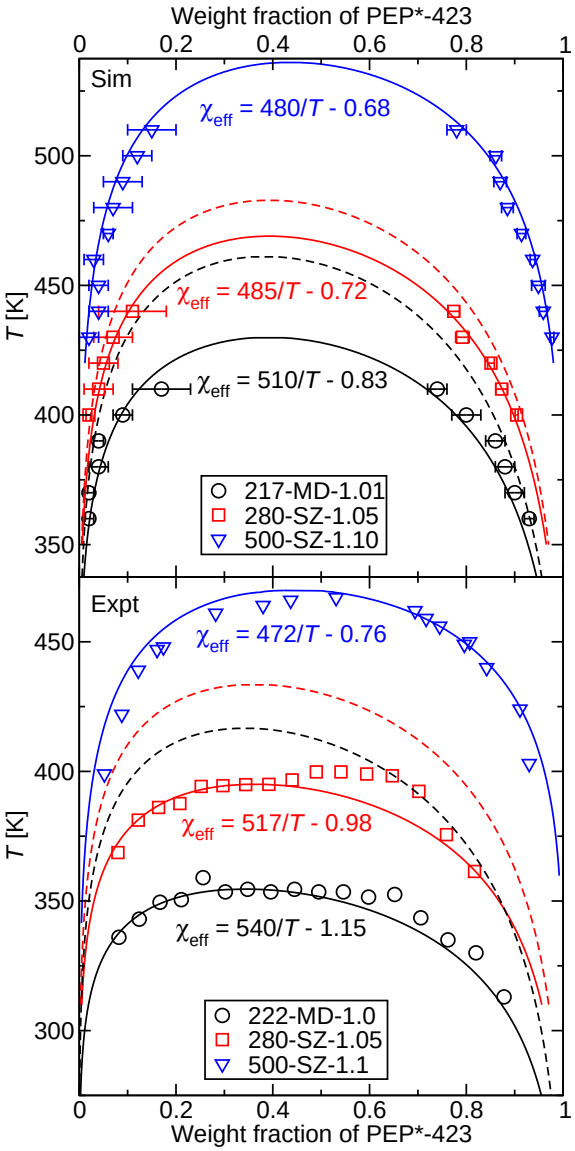

Figure 3.6: Simulated coexistence curves (top) and experimental cloud point curves (bottom) of PEP*-423/PEO mixtures. The experimental data for PEO-500-SZ-1.1 are from Washburn et al. [157], and the other two data sets were provided by courtesy of Shuyi Xie and Reza Foudazi. Solid lines are the fitting results using eq 3.1 with $\chi_{\text{eff}}(M,T) = a(M)/T + b(M)$, in which a and b are fitting parameters. Dashed lines are calculated phase diagrams using eq 3.1 with molecular weight independent $\chi_{\text{eff}}(T)$ from the PEO-500 fit. Analytical expressions of the molecular weight-dependent fits are provided, and the uncertainties of the fitting are estimated to be $\Delta\chi = \pm 0.01$, equivalent to UCST $= \pm 5$ K.

UCSTs of PEP*-423/PEO-222(PEO-217) and PEP*-423/PEO-280 mixtures are overestimated by about 30 and 60 K, respectively, for both simulation and experimental data sets. Therefore, one can conclude that χ of the PEP/PEO mixtures exhibits a non-negligible molecular weight dependence, and that this dependence stands in contrast to the general trend in earlier literature (decreasing χ with increasing the PEO molecular weight [31, 124, 158, 48]).

VLE of oligomeric PEP/PEO mixtures. To further characterize and rationalize the molecular weight dependence of χ, we performed VLE simulations to compute χ_{CP} for oligomeric PEP/PEO mixtures ($\phi_{\mathrm{PEP}} \approx 0.5$) with n ranging from two to six, and results are displayed in Figure 3.7. These simulations reproduce the PEO molecular weight dependence of χ_{eff} from LLE simulations. For both PEP-142 and PEP-212, increasing the molecular weight of PEO results in an increase in χ_{CP}. In contrast, χ_{CP} decreases with the increasing molecular weight of PEP, which displays the molecular weight dependence found in other experimental studies [28, 31]. More importantly, molecular weight exerts a large impact on χ_{CP} at both temperatures investigated. For example, varying the PEP chain length can result in a difference in χ_{CP} by as much as a factor of two (or 0.2 in absolute magnitude) for the PEP/PEO-222 mixture at $T = 490$ K. This corresponds to a difference in the coexistence curve by about 100 K. Therefore, it is crucial to understand different molecular weight dependences for the two components and to develop an approach to predict χ for PEP/PEO pairs of arbitrary chain lengths.

The Hildebrand formalism states that χ is proportional to the solubility parameter differences between two components:

$$\chi_{\mathrm{H}}(T) = \frac{v_{\mathrm{ref}}}{k_{\mathrm{B}}T}[\delta_1(T) - \delta_2(T)]^2 = \frac{v_{\mathrm{ref}}}{k_{\mathrm{B}}T}\left[\sqrt{\Pi_{\mathrm{CED},1}(T)} - \sqrt{\Pi_{\mathrm{CED},2}(T)}\right]^2 \tag{3.3}$$

where v_{ref} is the reference volume. The above equation can be used to qualitatively

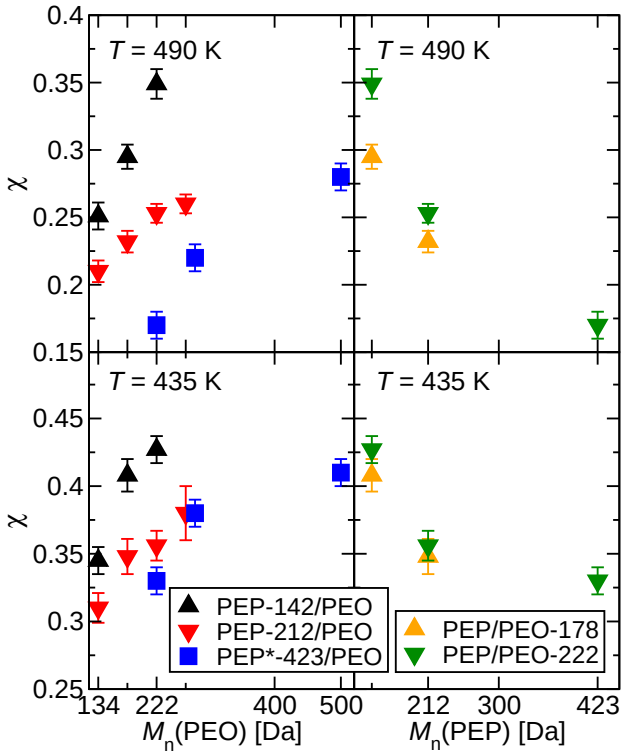

Figure 3.7: χ of oligomeric PEP/PEO mixtures as a function of the number average molecular weight of PEP or PEO at $T = 435$ and 490 K. Data that contain PEP*-423 are χ_{eff} from LLE simulations (see Figure 3.6), while the remainder of the data are χ_{CP} from VLE simulations.

explain the molecular weight dependence of χ. Figure 3.8a shows the cohesive energy density Π_{CED} as a function of chain length that was used to compute χ from the Hildebrand formalism. One can see that Π_{CED}(PEO) is always higher than Π_{CED}(PEP) at both temperatures (e.g., Π_{CED}(PEO-222) is 1.7 times higher than that of Π_{CED}(PEP-212) at $T = 435$ K, despite the similar molecular weights). As the molecular weight of PEP increases, $[\delta(\text{PEP}) - \delta(\text{PEO})]^2$ shrinks, and thus, so does χ. On the other hand, when the molecular weight of PEO increases, $[\delta(\text{PEP}) - \delta(\text{PEO})]^2$ becomes larger, which leads to the increase in χ. This qualitative agreement suggests that the molecular weight

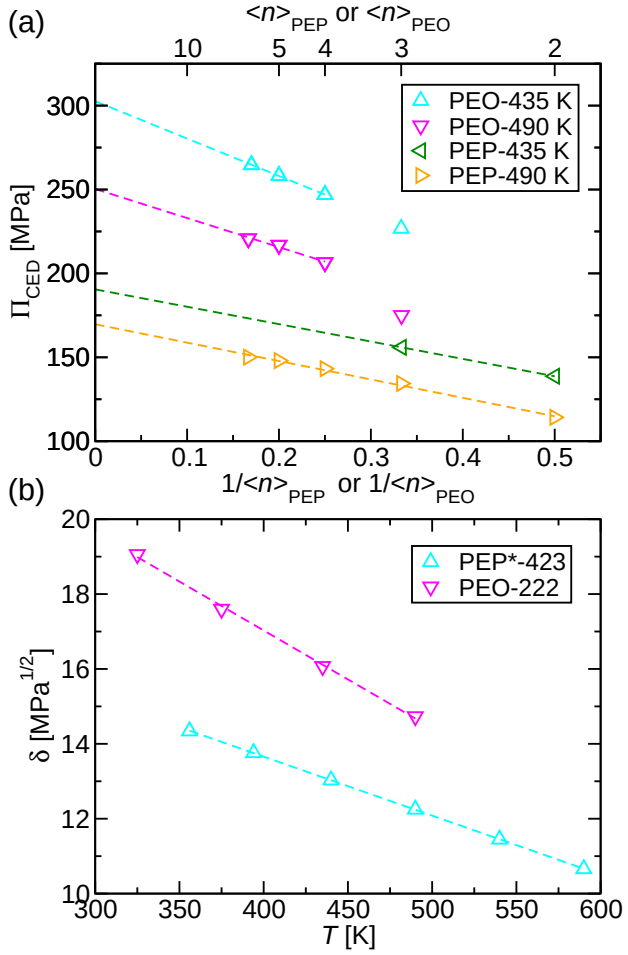

Figure 3.8: (a) Cohesive energy density, Π_{CED}, as a function of the inverse chain length of PEP and PEO. Dashed lines are linear fits to the data using data points with number of heavy atoms in PEP or PEO greater than or equal to 10 ($n_{\mathrm{PEP}} \geq 2$ and $n_{\mathrm{PEO}} \geq 4$). (b) Solubility parameter, δ, as a function of temperature for PEP*-423 and PEO-222. Dashed lines are linear fits of the data.

dependence of δ is the origin of the molecular weight dependence of χ, which implies that for such polar/nonpolar polymeric mixtures, the sign of the molecular weight dependence of χ is always opposite for the two components. For other mixtures such as

polyolefins, this argument may not hold due to the relatively greater effect of other non-idealities on the mixing thermodynamics (*e.g.,* finite compressibility, nonideal entropy of mixing) [125].

To quantify the molecular weight dependence of δ, two linear relationships were utilized. First, Figure 3.8a reveals that Π_{CED} is linear with the inverse chain length when oligomers are sufficiently long (when the number of heavy atoms is greater than or equal to 10), which enables the estimation of $\Pi_{\mathrm{CED}}(\mathrm{PEP})$ and $\Pi_{\mathrm{CED}}(\mathrm{PEO})$ of arbitrary chain length at $T = 435$ and 490 K. Second, as shown in Figure 3.8b, δ is found to vary linearly with temperature for temperatures well below the vapor–liquid critical temperature. This further allows for the extrapolation of Π_{CED} and δ to other temperatures. The following equations were extracted from the two linear fits and can be used to obtain Π_{CED} and δ for PEP and PEO of arbitrary chain length and temperatures.

$$\Pi_{\mathrm{CED}}(\mathrm{PEP}, 435\mathrm{K})/\mathrm{MPa} = -(103.6 \pm 0.3)/n + (190.5 \pm 0.4) \quad (n \geq 2) \tag{3.4}$$

$$\Pi_{\mathrm{CED}}(\mathrm{PEP}, 490\mathrm{K})/\mathrm{MPa} = -(109.6 \pm 0.3)/n + (169.7 \pm 0.5) \quad (n \geq 2) \tag{3.5}$$

$$\Pi_{\mathrm{CED}}(\mathrm{PEO}, 435\mathrm{K})/\mathrm{MPa} = -(222.7 \pm 0.2)/n + (302.7 \pm 0.3) \quad (n \geq 4) \tag{3.6}$$

$$\Pi_{\mathrm{CED}}(\mathrm{PEO}, 490\mathrm{K})/\mathrm{MPa} = -(173.4 \pm 0.3)/n + (250.3 \pm 0.4) \quad (n \geq 4) \tag{3.7}$$

$$\delta/\mathrm{MPa}^{1/2} = \left[\sqrt{\Pi_{\mathrm{CED}}(490\mathrm{K})} - \sqrt{\Pi_{\mathrm{CED}}(435\mathrm{K})}\right]\frac{T - 435}{55} + \sqrt{\Pi_{\mathrm{CED}}(435\mathrm{K})} \tag{3.8}$$

A more quantitative model is constructed by using the following equation [159, 88], in which k_{ij} is the only adjustable parameter that quantifies the deviation from the Berthelot combining rule used in the Hildebrand formalism:

$$\chi_{\mathrm{SP}}(T) = \frac{v_{\mathrm{ref}}}{k_{\mathrm{B}}T}\left[\Pi_{\mathrm{CED},1}(T) + \Pi_{\mathrm{CED},2}(T) - 2k_{ij}\sqrt{\Pi_{\mathrm{CED},1}(T)\Pi_{\mathrm{CED},2}(T)}\right] \tag{3.9}$$

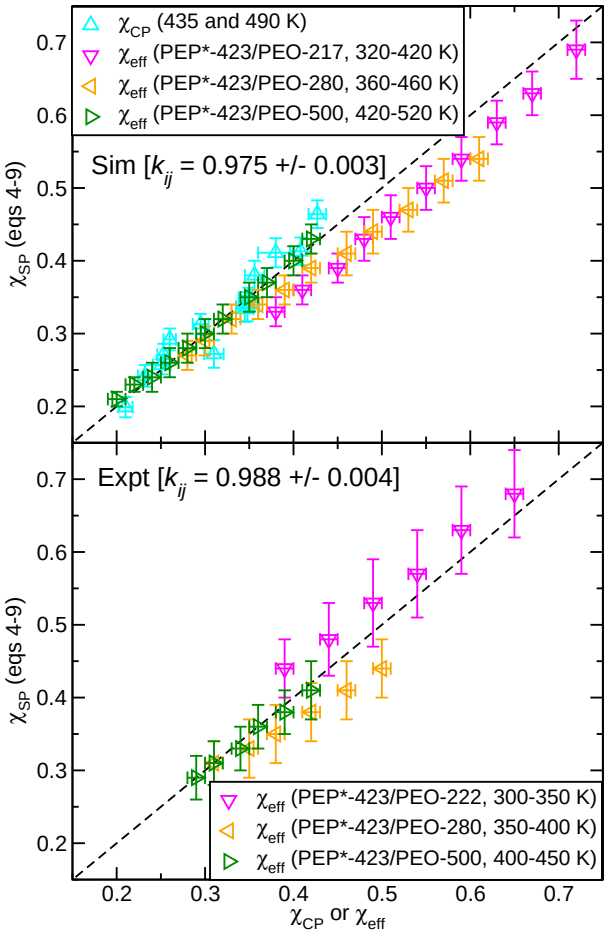

Figure 3.9: χ_{SP} as a function of χ_{CP} or χ_{eff} for PEP/PEO mixtures. k_{ij} values are fitted using χ_{CP} (cyan) and all data points for simulation and experiments, respectively.

When k_{ij} is unity, eq 3.3 is recovered. For simulations, empirical fitting to the χ_{CP} values yields a k_{ij} value of 0.975 ± 0.003. The utility of this k_{ij} value is confirmed by using χ_{eff} from LLE simulations as the validation set. Figure 3.9 reveals that the calculated χ_{SP} using this k_{ij} value agrees with χ_{CP} and χ_{eff}, with MUPE for all the data points being $6 \pm 4\%$ (6% is roughly 0.02 in absolute magnitude) and the maximum absolute error being 0.06 ± 0.03. The average error approaches the average uncertainty of the simulation data. It corresponds to a deviation in the mixture UCST of about 10 K, while the maximum

error corresponds to a deviation of about 25 K. As for experiments, due to the lack of χ for shorter PEP/PEO mixtures, all the χ_{eff} data are used for the fitting, which results in a k_{ij} value of 0.988 ± 0.004. The MUPE for all the experimental data points are $6 \pm 4\%$, and the maximum absolute error is 0.06 ± 0.04. The discrepancy in simulated and experimental k_{ij} values reflects the inaccuracy of the PEP-PEO interactions in the non-polarizable molecular model. If the force field were sufficiently accurate, k_{ij}(sim) and k_{ij}(expt) should converge, which implies that simulations can accurately predict $\chi(M,T)$.

To demonstrate that χ_{SP} computed from eq 3.9 can be used to calculate coexistence curves, predicted coexistence curves using the fitted k_{ij} values as well as a k_{ij} value of unity are shown in Figure 3.10. One can see that the maximum error between simulations/experiments and eq 3.9 predictions using $k_{ij}(\text{sim}) = 0.975/k_{ij}(\text{expt}) = 0.988$ is less than 20 K. When a k_{ij} value of 1.0 is used, the predicted coexistence curves deviate significantly from the simulation or experimental data, with the average error greater than 100 K. This result suggests that it is essential to fit a k_{ij} value from existing simulation or experimental data in order to predict χ and coexistence curves of binary polymer blends with arbitrary molecular weights.

The data in Figure 3.10 show large deviations for the mixtures with the two shorter PEO chains that exhibit miscibility gaps shifted to lower temperatures. This may be an indication of a weak temperature dependence for k_{ij}, which is consistent with recent calculations of water/oil interfacial tension [159]. The explanation for why k_{ij} for this system is smaller than unity is that PEO-PEO interactions include first-order electrostatic contributions that are not present for the PEP-PEP and PEP-PEO interactions (lumping first- and second-order electrostatic interactions). Therefore, PEP-PEO interactions predicted from eq 3.3 are too favorable, and thus, a k_{ij} value that is smaller than one is essential to correct the underestimation in χ. Similarly, k_{ij}(sim) is smaller than

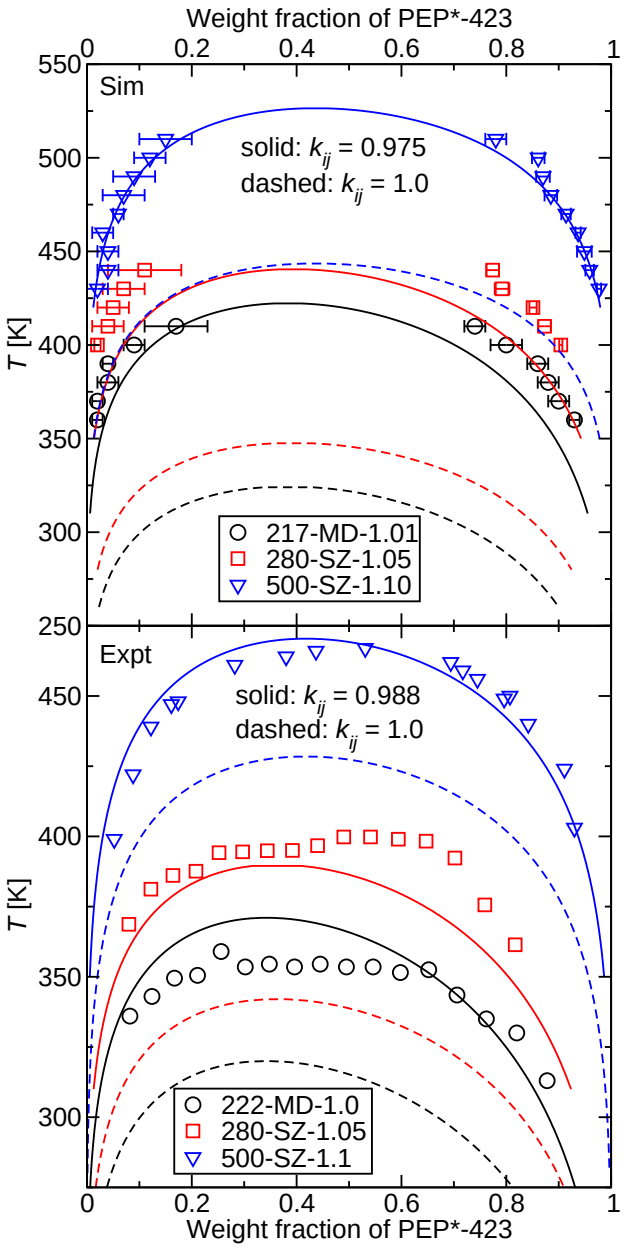

Figure 3.10: Predicted coexistence curves for PEP*-423/PEO mixtures for simulations (top) and experiments (bottom). The symbols represent simulation and experimental (by Shuyi Xie, Reza Foundazi, and Newell Washburn) data. Solid lines are from eq 3.9 using k_{ij} values of 0.975 and 0.988 for simulation and experiments, respectively. Dashed lines are from eq 3.9 using a k_{ij} value of 1.0.

k_{ij}(expt) because the dipole–induced dipole interactions that are present in real systems are missing in the non-polarizable TraPPE-UA model, leading to the overestimation in χ for the simulated systems. Furthermore, the origin of the temperature dependence of k_{ij} is the temperature dependent contribution from first-order dipole-dipole interactions due to the Boltzmann averaging of orientation [97]. However, despite the success in the qualitative agreement mentioned above, we know from the binodal curve fitting that an entropic deviation from the FH theory is also lumped into χ_{eff}. Future studies on other molecular systems with wider temperature ranges and calculation/measurement of the heat of mixing are needed to elucidate the physical meaning of k_{ij}.

3.3.2 The effects of MWD and Đ on the phase diagram

LLE of PEP*-423/PEO-500 mixtures. Figure 3.11 displays liquid–liquid coexistence curves from simulations and experimental cloud point curves (by courtesy of Shuyi Xie) for quasi-binary mixtures consisting of near-monodisperse PEP*-423 (Đ = 1.04) and disperse PEO with $M_{\text{n}} \approx 500$ Da, but with various MWDs and Đ values. Simulated coexistence curves corresponding to PEO-500-SZ-1.02, PEO-500-SZ-1.10, PEO-500-SZ-1.19, and PEO-500-BM-1.12 (Đ varying from 1.02 to 1.19) are indistinguishable from each other within statistical uncertainties. This is confirmed by the near identical experimental cloud point curves for PEO-500-SZ-1.1 and PEO-500-BM-1.1. These results reveal that regardless of MWD, Đ exerts negligible influence on the phase diagram when Đ is smaller than 1.2. Similar conclusions have also been reached by other experimental studies for high molecular weight polymers in solution [127].

However, coexistence curves of PEP*-423/PEO-500-BM-1.76, in which the PEO has the "extreme" bimodal distribution containing 57 mol% (15 wt%) of PEO-134 and 43 mol% (85 wt%) of PEO-1015 (Đ = 1.76), show significant deviation from those with smaller dispersities. The binary representation of the phase diagram is shown in Figure

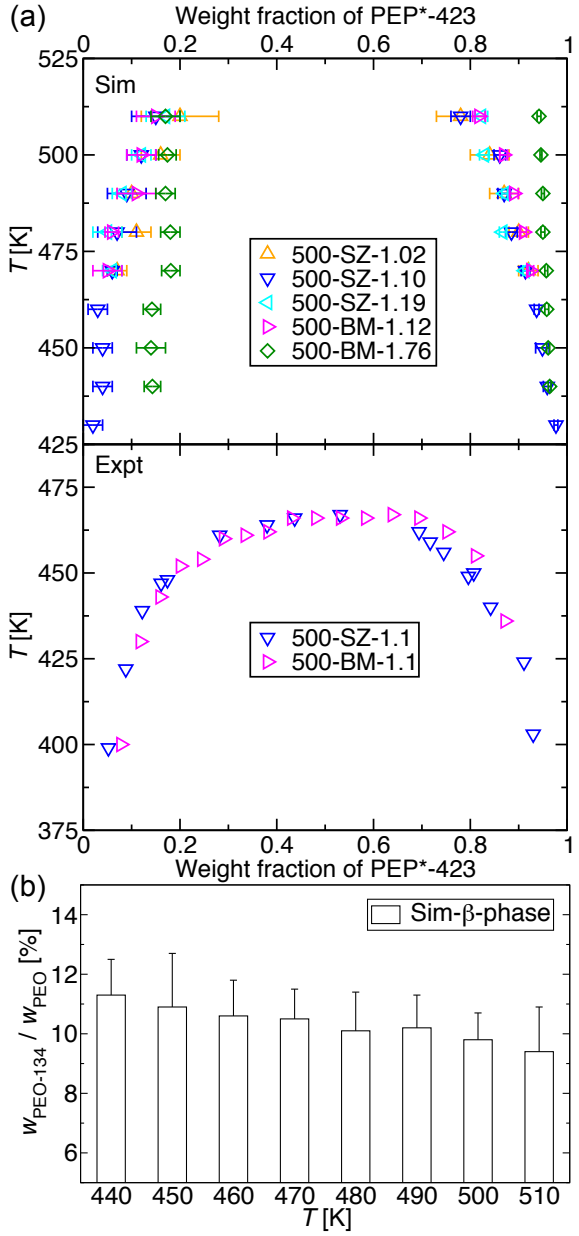

Figure 3.11: (a) Simulated coexistence curves (top) and experimental cloud point curves (bottom) for PEP*-423/PEO-500 mixtures. Experimental data for PEO-500-SZ-1.1 are from Washburn et al. [157], and the other experimental data set was provided by courtesy of Shuyi Xie. (b) Weight fraction of PEO-134 among all PEO molecules in the PEO-rich phase for the PEP*-423/PEO-500-BM-1.76 mixture.

3.11a and the ternary phase composition data are provided in Figure 3.11b and Table 3.3. Despite the increased complexity of the ternary phase diagram, the two major assumptions in the simulations are still valid. First, the liquid–liquid equilibria setup is appropriate because the molecular weight ratio between the two PEO components $r = M(\text{PEO-1015})/M(\text{PEO-134})$ is 7.6, which is lower than the three-phase separation limit of $r^* = 13$ from the generalized FH theory [160]. Furthermore, we cannot transfer PEO-1015 to the alkane phase due to the lack of medium molecular weight PEO molecules as transfer intermediates. Nevertheless, the assumption that PEO-1015 always stays in the PEO-rich phase is supported by the negligible solubility of PEO-1015 in PEP*-423 (less than 1 wt% at $T = 500$ K), estimated from eq 3.1 with a χ_{SP} value from eqs 3.4–3.9.

Table 3.3: LLE data for the PEP*-423/PEO-500-BM-1.76 mixture from simulations. Data are shown in the weight (w) and mole (x) fraction of PEP*-423 (1) and PEO-134 (2).

T [K]	w_1^α	w_2^α	w_1^β	w_2^β	x_1^α	x_2^α	x_1^β	x_2^β
440	0.964_3	0.036_3	0.143_{17}	0.097_{11}	0.899_9	0.101_9	0.20_2	0.388_{10}
450	0.961_3	0.039_3	0.14_3	0.093_{17}	0.892_7	0.108_7	0.20_3	0.378_{11}
460	0.958_4	0.042_4	0.142_{18}	0.091_{12}	0.887_{10}	0.113_{10}	0.20_2	0.371_{13}
470	0.957_3	0.043_3	0.181_{19}	0.086_9	0.883_8	0.117_8	0.25_2	0.347_{13}
480	0.950_4	0.050_4	0.18_2	0.082_{11}	0.866_{10}	0.134_{10}	0.26_3	0.336_{17}
490	0.950_3	0.050_3	0.17_2	0.084_{10}	0.866_7	0.134_7	0.24_2	0.343_{11}
500	0.946_2	0.054_2	0.174_{18}	0.081_9	0.857_5	0.143_5	0.25_2	0.334_{10}
510	0.942_4	0.058_4	0.17_3	0.078_{13}	0.849_{10}	0.151_{10}	0.25_4	0.326_{18}

Two features in the coexistence curve of the PEP*-423/PEO-500-BM-1.76 mixture stand in contrast to those containing low-dispersity PEO. First, the quasi-binary compositions are relatively insensitive to temperature, with the solubility of PEP*-423 in the PEO-rich phase changing from 14.3 ± 1.7 wt% at $T = 440$ K to merely 17 ± 3 wt% at $T = 510$ K. This can also be deduced from the weight fraction of PEO-134 among all the PEO molecules in the PEO-rich phase (this fraction is unity in the PEP-rich phase), shown in

Figure 3.11b. The decrease of PEO-134 weight fraction as a function of temperature is on the order of the simulation uncertainties. The temperature insensitivity can mainly be attributed to the dominant overall weight fraction of PEO-1015 among all the PEO molecules (85 wt%). Consequently, the distribution change of PEO-134 hardly influences the coexistence curve. For example, when the temperature increases from 440 to 510 K, the mole fraction of PEO-134 in the PEP-rich phase increases by around 50% (from 10.1 ± 0.9 to 15.1 ± 1.0 mol%, or from 3.6 ± 0.3 to 5.8 ± 0.4 wt% on the weight fraction scale). The second feature of interest is that the PEP solubility in the PEO-rich phase is much higher at low temperatures ($T \leq 500$ K) compared to that of PEO-500-BM-1.12. For example, at $T = 470$ K, the PEP solubility in PEO-500-BM-1.76 (18.1 ± 1.9 wt%) is about a factor of three to four higher than that of PEO-500-BM-1.12 (5 ± 3 wt%). The increase in solubility can be explained by the more favorable interactions between PEP and PEO, which will be more quantitatively elaborated later when the PEP free energy of transfer data are discussed. This deviation from other data sets clearly shows the failure of treating this ternary mixture as a binary one. Unfortunately, the theoretical treatment of the ternary mixture can be much more complicated since the binary interaction parameters χ is a function of composition when the third component is the majority species in weight or volume fraction [161]. In summary, this result reveals that for mixtures with a bimodal distribution and a large dispersity, which can be the case for many industrial polymers, the treatment of the mixture as quasi-binary blend can introduce non-negligible errors.

The effect of dispersity on molecular partitioning raises the question of why the overall phase diagram is not significantly perturbed by changing from SZ-1.02 to SZ-1.19. Mixtures containing PEO with larger dispersity (e.g., PEO-500-SZ-1.76) would be of interest, but are more challenging for simulation because of the need for a larger number of components to represent the SZ versus the BM distribution. The molecular

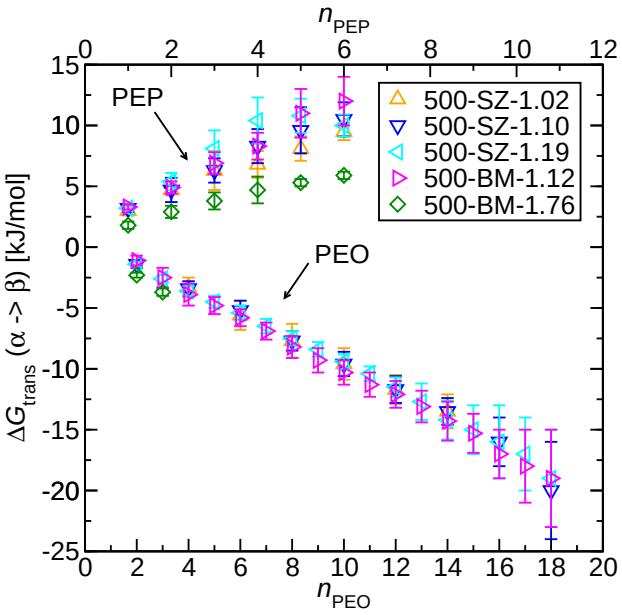

Figure 3.12: Free energy of transfer (ΔG_{trans}) from the PEP-rich (α) to the PEO-rich phase (β) as a function of the PEP and PEO chain lengths at $T = 470$ K.

weight dependence of the free energy of transfer ΔG_{trans} can be used to address this question, which is calculated from the following equation [162, 97, 102]:

$$\Delta G_{\text{trans}}(\beta \rightarrow \alpha) = -k_{\text{B}} T \ln(\rho_\alpha / \rho_\beta) \tag{3.10}$$

where ρ_i is the number density of PEP or PEO in phase i. Note that ρ_i can be obtained from either simulations or experiments using the composition and the density of phase i. Figure 3.12 depicts ΔG_{trans} from the PEP-rich (α) to the PEO-rich (β) phase as a function of the chain length for both PEP and PEO at $T = 470$ K. ΔG_{trans} is found to be a linear function of chain length for both types of molecule, and the values do not change with respect to the MWD or Đ, with the exception of PEO-500-BM-1.76. For the mixtures with Đ < 1.2, the incremental transfer free energies are found to be

2.0 ± 0.4 and 1.0 ± 0.1 kJ/mol for one PEP and PEO repeat unit, respectively.

One can obtain the effect of dispersity on compositions and MWDs for both phases using continuous thermodynamics [131, 132, 133]. For the partitioning of PEO between two phases, the following equations hold:

$$K_\alpha = \frac{\rho_\alpha}{\rho_\beta} = \exp\left[\frac{-\Delta G_{\text{trans}}(\beta \to \alpha)}{k_{\text{B}}T}\right] = \exp\left[\frac{-k_{\text{trans}}(\beta \to \alpha)M_{\text{PEO}}}{k_{\text{B}}T}\right] \tag{3.11}$$

in which K_α is the partition constant and $\Delta G_{\text{trans}}(\beta \to \alpha) = k_{\text{trans}}(\beta \to \alpha)M_{\text{PEO}}$. This equation shows that K increases or decreases exponentially as the chain length increases. The total mass of PEO partitioned in both phases, m_α^{PEO} and m_β^{PEO}, can be expressed by the following two equations:

$$m_\alpha^{\text{PEO}} = \int_0^\infty P_\alpha(M)dM = \int_0^\infty P(M)\frac{K_\alpha(M)K_v}{K_\alpha(M)K_v + 1}dM \tag{3.12}$$

$$m_\beta^{\text{PEO}} = \int_0^\infty P_\beta(M)dM = \int_0^\infty P(M)\frac{1}{K_\alpha(M)K_v + 1}dM \tag{3.13}$$

in which $P(M)$ is the total MWD in mole fraction; $P_i(M)$ is the MWD in phase i; K_v is the volume ratio between the two phases, and can be calculated via:

$$K_v = \frac{V_\alpha}{V_\beta} = \frac{m_\alpha/\rho_\alpha}{m_\beta/\rho_\beta} = \frac{\rho_\beta}{\rho_\alpha}\frac{(w_\beta - w_{\text{tot}})}{(w_{\text{tot}} - w_\alpha)} \tag{3.14}$$

where V_i, m_i, ρ_i, and w_i are the volume, mass, density, and PEO weight fraction of phase i, respectively; w_{tot} is the total weight fraction of PEO. If one assumes the partitioning of PEP*-423 is not affected by the dispersity of PEO (free energy of transfer of PEP is not a function of dispersity from Figure 3.12), eqs 3.11–3.14 can be solved together with the mass balance to compute w_i and K_v given the total weight fraction w_{tot} and the original MWD $P(M)$. The solution can provide the effect of dispersity on the coexistence curve

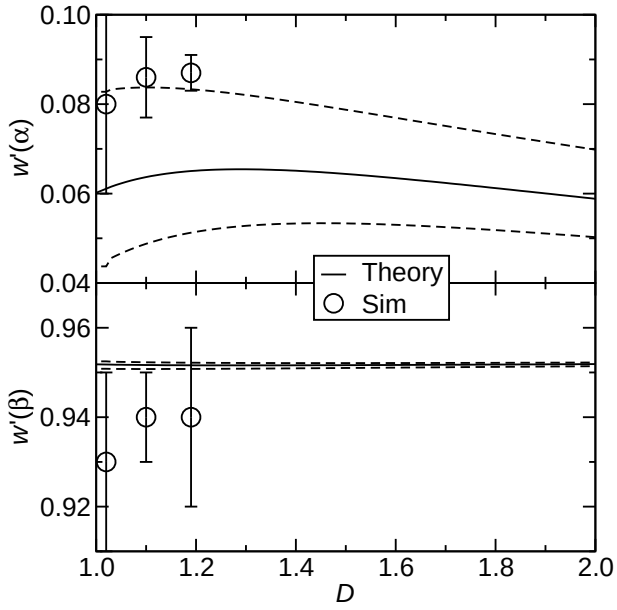

Figure 3.13: The weight fraction of PEO in α (top) and β (bottom) phase as a function of dispersity for the PEP*-423/PEO-500 mixture when PEO has the Schulz-Zimm overall distribution at $T = 470$ K. Symbols are simulation data, while solid and dashed lines are the calculated values and uncertainties from eqs 3.11–3.14 and Figure 3.12, respectively.

and quantify the MWD in each phase.

Figure 3.13 shows the calculated PEO compositions in both phases as a function of dispersity for the PEP*-423/PEO-500 mixture when PEO has a Schulz-Zimm overall distribution. The $k_{\text{trans}}(\beta \to \alpha)$ value of 22.5 ± 2.5 J/g from Figure 3.12 is used. One can see that the calculated compositions agree fairly well with the simulated ones within uncertainties, which serves as a self-consistency check. In addition, the lines are almost flat, which suggests that dispersity does not exert a large influence on the coexisting compositions for the SZ distribution at $T = 470$ K for $Đ \leq 2$. The result is consistent with the simulation and experimental findings for PEP*-423/PEO-500-SZ mixtures. The comparison between the predicted coexistence curve for PEP*-423/PEO-500-SZ-1.76 and the simulated coexistence curve for PEP*-423/PEO-500-BM-1.76 suggests that the

shape of the MWD plays a pivotal role in the effect of dispersity on the polymer mixture phase diagram (e.g., the effect of dispersity is more obvious for the BM MWD). For other overall MWDs and molecular weights, one can use the numerical method formulated here to obtain an estimate regarding the effect of dispersity.

Theoretical calculation reveals that the change in free energy of transfer is the key to the shift in composition for the PEP*-423/PEO-500-BM-1.76 mixture. From Figure 3.12, one can see that the free energy of transfer of PEP in the PEP*-423/PEO-500-BM-1.76 mixture at $T = 470$ K deviates substantially from the rest of the data set. The smaller free energy barrier for PEP to transfer to the PEO-rich phase is the origin of the higher PEP concentration in the PEO-rich phase. Similarly, the slightly higher free energy barrier for low molecular weight PEO to transfer into the PEP-rich phase coincides with the lower PEO-134 composition in the PEP-rich phase at equilibrium for the PEP*-423/PEO-500-BM-1.76 mixture. The shift in transfer free energy implies that the quasi-binary approximation fails when the dispersity of PEO is as large as 1.76. It also suggests that the binary interaction parameter χ is a function of the composition in the ternary mixture. For example, when the composition of the PEO-rich phase changes from PEO-500-SZ-1.19 to PEO-500-BM-1.76 (the compositions of the PEP-rich phase are similar), it becomes more favorable for PEO-134 to transfer into the PEO-rich phase, with ΔG_{trans} changing from -2.6 ± 0.3 to -3.7 ± 0.3 kJ/mol, due likely to the presence of high-molecular weight PEO molecules.

Molecular partitioning. The partitioning of different molecular weight PEO molecules is revealed by the MWD in each phase from LLE simulations (symbols) and theoretical calculations (lines), as shown in Figure 3.14. The agreement between simulation data and theoretical calculations demonstrates the robustness of the numerical formulation (eqs 3.11–3.14). When the dispersity is low ($Đ = 1.02$), the MWDs for both phases are almost identical. When the dispersity increases to $Đ = 1.10$, a distinction between the

two distributions emerges at the higher molecular weight end of the distribution. For example, the mole fraction of PEO-707 among all the PEO molecules in the PEO-rich phase is 80% higher than that in the PEP-rich phase. This indicates an enrichment of higher molecular weight PEO molecules in the PEO-rich phase, which is consistent with the lower solubility of higher-molecular weight PEO in the PEP-rich phase. When Đ further increases to 1.19, the distinction between two MWDs becomes more pronounced.

This effect of MWD can also be quantified by the M_{n}(PEO) present in each phase shown in Figure 3.14b. It is apparent that for both types of MWD, M_{n}(PEO) is higher in the PEO-rich phase, and the gap between the two M_{n} values in each phase grows, as the overall dispersity increases. When Đ = 1.19, M_{n}(PEO) in the PEO-rich phase is 50% higher than that in the PEP-rich phase at $T = 470$ K. For the bimodal distribution, M_{n}(PEO) in both phases are lower than those of the Schulz-Zimm distribution when Đ values are close. This is due to the higher abundance of lower molecular weight PEO in the overall MWD as well as fewer high molecular weight molecules in the Schulz-Zimm distribution. The difference in partitioning is the thermodynamic driving force of fractionation of disperse polymers [140, 163]. Moreover, Figure 3.14 also reveals that the MWDs of polymers in each phase can differ even if the overall compositions in the phase diagram are similar. For example, when $T = 470$ K, M_{n}(PEO) in the PEP-rich phase can vary from 462 ± 7 Da to 352 ± 6 Da when Đ increases from 1.02 to 1.19, but the corresponding phase diagrams are nearly identical (see Figure 3.11). Therefore, the approximation that such a quasi-binary mixture can be treated as a two-component system can be misleading, and caution needs to be taken if the exact content of the polymer in each phase is of interest.

Structural analysis. The effect of liquid structure on the mixing thermodynamics is characterized by the intermolecular radial distribution functions (RDFs) computed from the simulation trajectories, plotted in Figure 3.15. For both PEP*-423/PEO-500-SZ-1.10

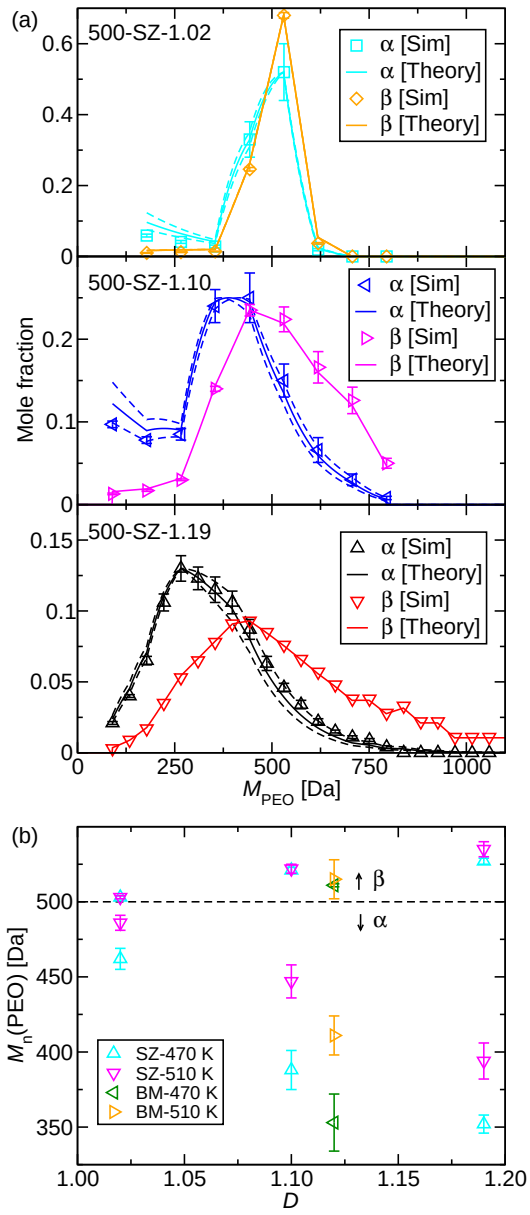

Figure 3.14: (a) PEO MWD in the PEP-rich (α) and the PEO-rich (β) phases at $T = 470$ K for PEP*-423/PEO mixtures with various dispersities. Symbols represent simulation data, while solid and dashed lines represent theoretical calculations and the uncertainties using eqs 3.11–3.14, respectively. (b) M_{n}(PEO) as a function of dispersity.

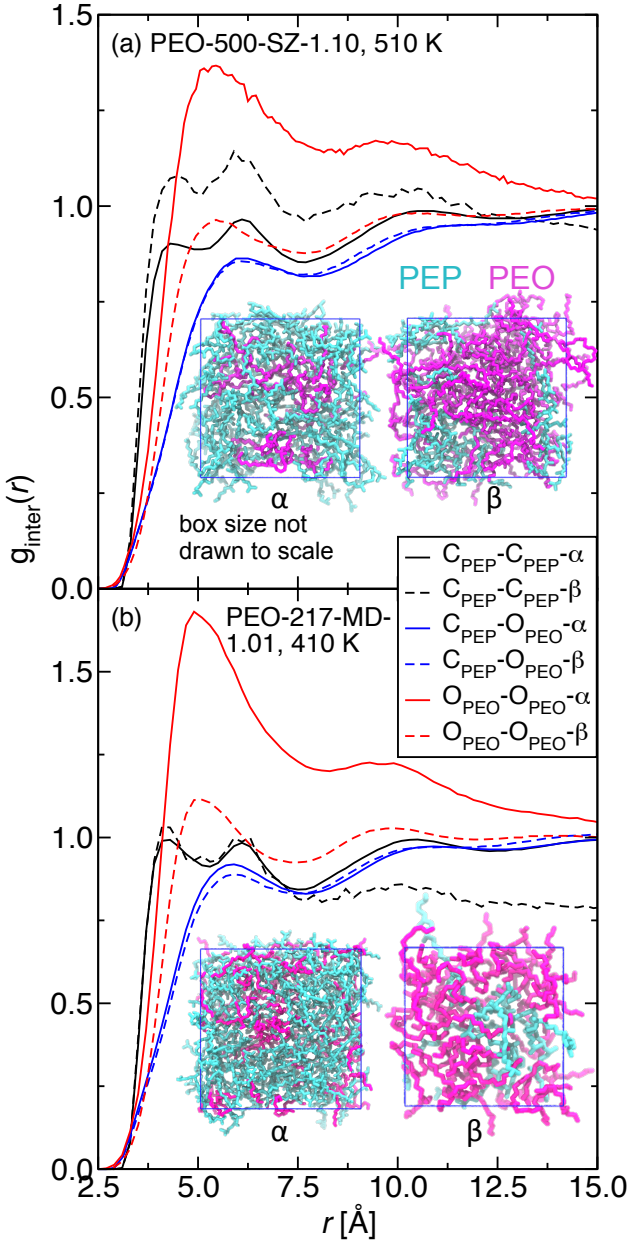

Figure 3.15: Intermolecular radial distribution functions $g_{\text{inter}}(r)$ and representative simulation snapshots for the two PEP*-423/PEO mixtures.

and PEP*-423/PEO-217-MD-1.01 mixtures at a relatively high temperature near their UCSTs, the peak positions of the oxygen–oxygen RDF for intermolecular PEO–PEO pairs in both phases (solid and dashed lines in red) are similar. Nevertheless, the height of the RDFs is vastly reduced in the PEO-rich phase compared to that corresponding to the PEP-rich phase. This is an indication of the clustering behavior of PEO molecules in the minority phase. Similar aggregation behavior can also be seen for PEP–PEP pairs in the PEO-rich phase containing PEO-500-SZ-1.10 (dashed black lines in the top panel), but the distinction between two carbon–carbon RDFs for the mixture with shorter PEO chains PEO-217-MD-1.01 is less obvious (solid and dashed lines in black in the bottom panel). In contrast, the carbon–oxygen RDFs for the PEP–PEO pair in two phases (solid and dashed lines in blue) are found to trace closely to each other, revealing no sign of PEP–PEO aggregation. The clustering behavior agrees with our finding that the entropic contribution to χ is always a significant portion of the total χ (i.e., the intercept b in the $\chi = a/T + b$ fitting is non-zero). In addition, the aggregation behavior is consistent with other simulations of polymer blends [34, 51] and block polymers [71, 164, 165, 69]. The simulation snapshots depicted in Figure 3.15 also confirm the clustering behavior. Both PEP and PEO molecules tend to aggregate when they are the minority components. This clustering behavior reflects the incompatibility of the two components and further reveals the complexity of mixing behavior that deviates significantly from the ideal mixing assumption of the FH theory.

3.4 Conclusions

This work uses GEMC simulation in conjunction with the experimental cloud point measurements (by courtesy of Shuyi Xie and Reza Foudazi) to study the phase behavior of oligomeric PEP/PEO mixtures. The utilization of a modified TraPPE–UA force

field for PEO greatly improves the accuracy of the computed χ_{CP} for alkane/ether mixtures, but it still results in an overestimation of χ_{CP} by around 35% (or 0.1 in absolute magnitude), and thus, overestimation of the coexistence curves for binary alkane/ether mixtures. Results from simulations shed light on the molecular weight dependence of χ and the effect of MWD and Đ on the liquid–liquid phase diagram. First, χ_{eff} or χ_{CP} of PEP/PEO mixtures increases as the molecular weight of PEO increases or as the PEP molecular weight decreases. This seemingly irregular dependence is rationalized by the chain length dependence of the cohesive energy densities for PEP and PEO. A modified Berthelot mixing rule with a single adjustable parameter k_{ij} has been established to rationalize the dependence and accurately predict χ for PEP/PEO mixtures of arbitrary molecular weights. For simulation data, k_{ij} is fitted to χ_{CP} of shorter oligomer mixtures and yields a value of 0.975 ± 0.003. This value differs from the k_{ij} fitted to the experimental cloud point curves of PEP*-423/PEO mixtures (0.988 ± 0.004) due to the inaccuracy of the force field. For both simulation and experiments, the predicted coexistence curves using these k_{ij} values agree with those for PEP*-423/PEO mixtures with less than 20 K error, while the deviation is more than 100 K if k_{ij} is unity (the Hildebrand formalism). Second, the binodal curves of PEP*-423/PEO-500 mixtures with varying MWD and Đ of PEO are investigated. Experiments and simulations both indicate no significant change in the coexisting phase compositions when the dispersity is below 1.2. However, the phase diagram shifts dramatically when a bimodal distribution is used (Đ = 1.76). The shift suggests that the quasi-binary approximation fails when one component has a bimodal distribution with a large Đ, and the ternary treatment is essential to describe the phase behavior of the mixture. The molecular partitioning extracted from simulations reveals the enrichment of high molecular weight PEO in the PEO-rich phase, which causes a decrease of M_n(PEO) in the PEP-rich phase as Đ increases. This suggests that PEO compositions can be quite different even if the phase

diagrams in weight fraction units are similar. Furthermore, structural analysis from simulations shows signs of aggregation for PEP and PEO molecules in their minority phases, which leads to deviations from the ideal mixing behavior and brings more complexity to the prediction of its phase behavior.

Chapter 4

Computational Design of High-χ Block Oligomers for Accessing 1-nm Domains

4.1 Introduction

In the microelectronics industry, the growing need for microprocessors with faster speed and storage devices with larger storage densities demands the miniaturization of the associated features below the current practical limit of about 10 nanometers [22]. Further feature size decreases using traditional lithographic processes faces significant challenges [166]. Nanolithography utilizing directed self-assembly of block polymers is considered an attractive and cost-effective approach to achieve the desired single digit nanometer length scales [167, 22]. For block polymers, the morphology and the separation between

This chapter is reproduced in part with permission from (Q. P. Chen, L. Barreda, L. E. Oquendo, M. A. Hillmyer, T. P. Lodge and J. I. Siepmann, "Computational Design of High-χ Block Oligomers for Accessing 1 nm Domains", *ACS Nano* **2018**, DOI: 10.1021/acsnano.7b09122, American Chemical Society)

self-assembled domains can be tailored by tuning the block polymer composition (f_i), the overall degree of polymerization (N), and the interaction parameter between the blocks (χ). Pioneering research has been devoted to exploring the materials design space and identifying, synthesizing, and implementing new high-χ block polymers (oligomers) that can be readily used for such applications [168, 169, 170, 171, 172]. In addition, a variety of processing techniques including thin-film coating [173], defect removal [174, 175, 176], and domain alignment [177, 178, 179] have been developed to make block polymers more practically viable. However, typical domain periods d (d is defined as the domain spacing for lamellae or the 100 interplanar distance for hexagonally packed cylinders) for self-assembled block polymers ($d \approx$ 15–100 nm) are still higher than the desired periods ($d \leq$ 5 nm) for nanolithography [180, 167, 22] and other applications [20, 14, 13, 181]. Since d increases with N, while χN must be large enough to access the microphase separation region, the key to reducing d is to find "high χ–low N" block polymers (oligomers) which contain more incompatible blocks [22].

There are numerous examples in the literature aimed at practical approaches to such high-χ–low N systems in attempts to drive down d [182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205]. For example, Kennemur and co-workers [196] have demonstrated that domain sizes of sub-5 nm can be achieved without the incorporation of silicon, fluorine, or metal species (i.e., "all-organic" block polymers), the presence of which is undesirable for convenience of post-processing and sustainability [22, 206, 207]. In addition, the use of highly-hydroxylated blocks has recently been proved a promising method. In particular, Otsuka et al. utilized maltoheptaose to fabricate block polymers with d as small as 7.7 nm [200]. Nowak et al. designed a sugar-olefin conjugate that achieved the nanostructure formation with $d = 6.6 \pm 0.7$ nm [203]. Similarly, Jeong et al. utilized a solid-state transition to self-assemble block oligomers that contain multiple hydroxyl groups into 5.6-nm-spaced

lamellae [204].

Further reduction in domain separation calls for predictive design of high-χ block oligomers (HCBOs). This, in turn, creates a demand for more comprehensive understanding of the phase behavior at this length scale, which lies at the boundary between self-assembled block polymers and solvent-free surfactants [22]. The microphase separation of block polymers is driven by the interplay between segment incompatibility and the enforced block connectivity. Thus, the associated complex phase behavior can be largely attributed to the chain flexibility that can induce preferred curvature at the interface, while the precise molecular geometry (e.g., head group area, tail group volume) plays a less important role [169]. On the other hand, low molar mass surfactants typically adopt only layered smectic or nematic phases in their solvent-free liquid crystalline state and not other complex morphologies [208, 209]. This liquid crystalline-type phase behavior stems largely from anisotropic interactions such as hydrogen bonding and π-π interactions, where the self-assembly behavior is heavily influenced by molecular geometry [210]. While the general principles for both block polymer and liquid crystalline surfactant self-assembly are well-established, it is still a matter of discussion as to whether these theories are applicable to sub-5 nm HCBOs [169, 191]. Therefore, the complication resulting from this interplay motivates a more rigorous understanding of the phase behavior at the sub-5 nm length scale, and molecular simulation offers a promising approach to address this challenge.

Earlier works in simulating block polymers (in fact, polymers in general) often utilize coarse-grained force fields for which the "resolution" of the model is relatively low (e.g., one monomer is modeled as a spherical interaction site and its chemical structure is obscured). These pioneering studies provide valuable insight into the phase behavior of block polymers [211, 66, 212, 213, 71, 214, 215], but coarse-grained force fields are insufficient for the predictive design of new block oligomers, because the detailed molecular

features play a more important role at shorter length scales. Therefore, atomistic-level force fields with higher resolution are used to model HCBOs. This allows for more refined descriptions of molecular-level interactions, and higher transferability of the force field parameters compared to coarse-grained molecular models. As a consequence, a direct relationship between molecular structure and phase behavior can be established without any extraneous fitting parameters. This provides a convenient route to identify HCBOs that can self-assemble into structures with sub-5 nm domains, and to generate further insight into critical molecular features governing self-assembly at the block polymer/surfactant boundary.

Molecular dynamics (MD) simulations were performed in this work to design HCBOs that can self-assemble into various complex ordered mesophases with domain periods ranging from 2 to 4 nm. These amphiphilic molecules consist of hydrophilic polyol head groups and alkanes as hydrophobic tail(s). The molecular structures of these oligomers are shown in Figure 4.1a. It is noteworthy that simulated oligomers are a mixture of different stereoisomers since the tacticity is not preserved in the generation of initial molecular structures. The design space (chain length, volume fraction, chain architecture, and temperature) is explored *in silico* via molecular simulations. Morphologies of these oligomers are identified through both visualization and quantitative analysis. I also probe key factors that govern the self-assembly behavior, such as chain orientation and hydrogen bonding, from the analysis of simulation trajectories to understand the interplay between "surfactant factors" (e.g., head group interactions, chain flexibility, interfacial curvature) and "block polymer factors" (χ, N, and f). Predictions are provided for a series of novel HCBOs that exhibit exciting features for nanotechnological applications and elucidate pivotal factors that contribute to the mesophase formation at the nanometer length scale.

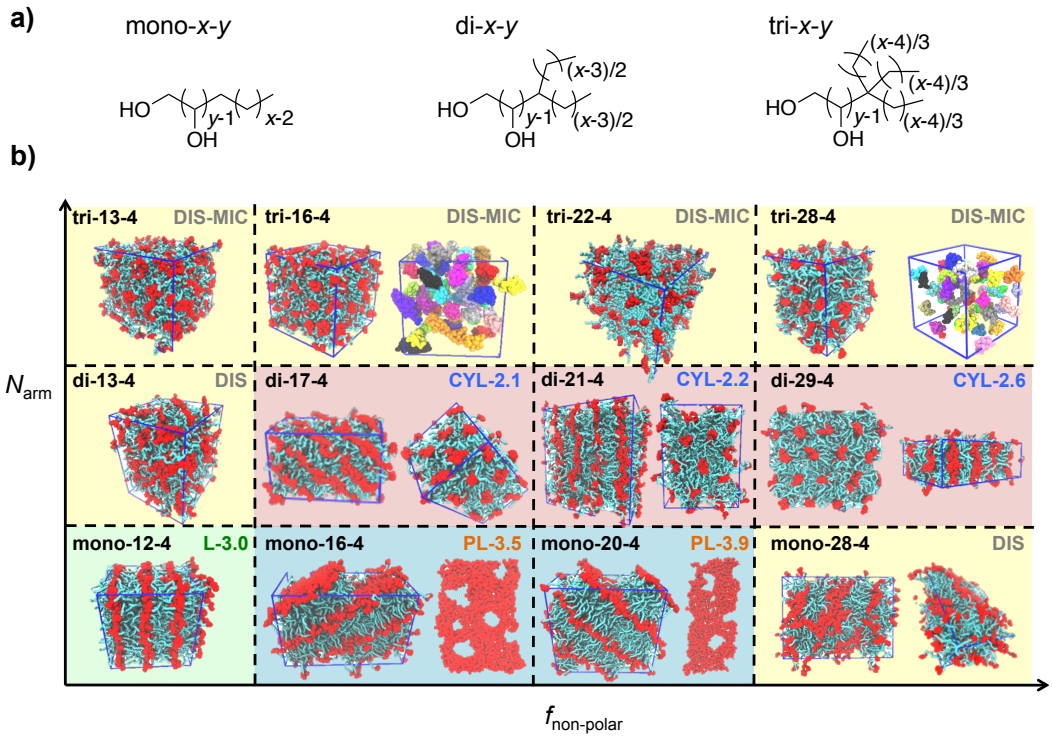

Figure 4.1: (a) Molecular structures of the HCBOs investigated. x and y indicate the number of hydroxyl groups and the number of CH_n groups in the alkyl tail(s), respectively. (b) Simulation snapshots of representative morphologies observed at $T_{\text{sim}} =$ 493 K. x and y indicate the number of hydroxyl groups and the number of CH_n groups in the alkyl tail(s), respectively. The caption labels the morphology followed by the domain period d (in nm). "L" = lamellae, "PL" = perforated lamellae, "CYL" = hexagonally-packed cylinders, and "DIS(-MIC)" = disordered(-micellar) phase. Atoms in the polar block are rendered in red (including the carbon atoms) and atoms in the non-polar block are rendered in cyan. For the two PL snapshots, one lamellae layer is depicted to show the holes. For both tri-16-4 and tri-28-4 snapshots, the second images show different micelles with different colors.

4.2 Simulation Methodology

4.2.1 Molecular models.

The oligomers studied in this work are represented by the Transferable Potentials for Phase Equilibria united-atom (TraPPE–UA) force field [95, 96, 216, 150]. The model represents the hydrogen and oxygen atoms of hydroxyl groups as individual interaction sites, and each CH_n (n = 0–3) group as one pseudo-atom located on the position of the carbon atom. The nonbonded interactions among interaction sites include pairwise-additive Lennard-Jones (LJ) and Coulomb potentials, which are calculated for all intermolecular pairs and for intramolecular pairs that are separated by four or more bonds. The potential is truncated at a spherical cutoff of 14 Å [98] and is shifted to ensure the continuity of the potential at the cutoff radius. Lorentz–Berthelot combining rules are used to describe the interactions between unlike interaction sites [97]. As implemented in the TraPPE–UA force field, a short range repulsive term between the oxygen and hydrogen atom that are separated by four bonds is used to prevent the unphysical overlap of the two sites (since the proton is not protected by the LJ potential and the 1–4 oxygen-oxygen LJ potential is not included). In the standard TraPPE–UA model, the interaction sites are joined by fixed bond lengths, but bending and dihedral angles are allowed to vary. In this work, the bond length is still fixed for bonds that contain protons to allow the use of larger time step, while harmonic stretching potentials are used for all the other bonds. The equilibrium bond length is taken from the TraPPE–UA force field and the stretching force constant is adopted from the OPLS force field [217].

4.2.2 Validation of molecular models.

The key to accurately predicting block oligomer phase behavior is to quantitatively reproduce the χ parameter between the two blocks, i.e., the interactions between the polyol

head group and the alkyl tail(s). Among the experimentally measurable quantities, the solubility of an alkane in a polyol is very sensitive to χ and, hence, it serves as the target property for the molecular model validation. In particular, the solubilities of propane in 1,2-ethanediol at various temperatures and pressures are calculated from Gibbs ensemble Monte Carlo simulations [92, 93] and compared to the experimental data [218], due to the unavailability of solubility data for longer alkane and polyols (likely because they are strongly immiscible).

Either a two-box vapor (or supercritical)–liquid or a three-box liquid–liquid–vapor (where the vapor-box functions as a transfer medium [62]) setup was used depending on whether the neat propane is a gas, supercritical fluid, or liquid at the state point of interest. The simulations were initiated with 400 1,2-ethanediol and 200 propane molecules in each simulation box. Translational, rotational, configurational, and volume moves were used to sample molecular configurations and the volume fluctuations [96]. In addition, propane and ethanediol were allowed to transfer between the simulation boxes [121, 62]. All the simulations were first equilibrated for 10^5 MC cycles (MCC, with one MCC consisting of N randomly chosen MC moves, where $N = 600$ is the number of molecules in the system), and then run for another 1.5×10^5 MCCs as the production period. Statistical uncertainties of the solubility data were estimated from 16 independent runs for each state point, and are reported as 95% confidence intervals. For all the simulations, a spherical truncation, r_{cut}, at 14 Å was used for the Lennard-Jones interactions unless otherwise indicated, and analytical tail corrections were used to account for the interactions beyond the cutoff distance [98]. The Ewald summation method [219] with parameters of $\kappa = 3.2/r_{\text{cut}}$ and $K_{\text{max}} = \kappa L_{\text{box}} + 1$ was used for the simulation box with Coulomb interactions.

Figure 4.2 summarizes the solubilities of propane in 1,2-ethanediol as a function of pressure at various temperatures from both simulation and experiments [218]. For both

sets, the solubility becomes higher as the pressure increases, and it saturates when the pressure is greater than 10 MPa (for $T_{\rm sim} < 400$ K). I focus on the saturated solubility because the two phases are more incompressible at higher pressures, and thus, represent better the binary interactions in the liquid phases. In general, solubility data from simulations are on average 60% lower than those from experiments at the same temperature (e.g., at $T_{\rm sim} = 348$ K, the deviations are -56, -69, -56, and -65% for $p = 0.1$, 2.85, 6.27 and 10.31 MPa, respectively.) This finding agrees with earlier simulation work [62] and implies that the χ parameter for alkane and polyols is overestimated.

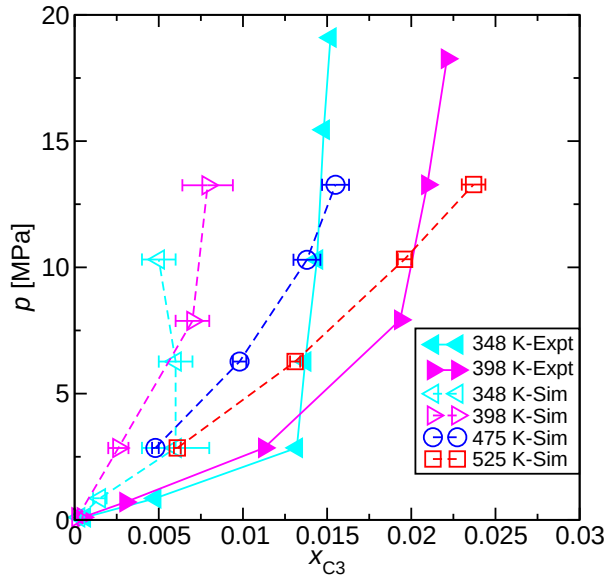

Figure 4.2: The solubility of propane in 1,2-ethanediol as a function of pressure. Experimental data are from Jou, Otto, and Mather [218].

To better quantify this deviation on the temperature scale, simulations at two elevated temperatures $T_{\rm sim} = 475$ and 525 K were also performed, and results are compared to the experimental solubilities at lower temperatures (deviations smaller than 5%). From Figure 4.2, the simulated solubilities at $T_{\rm sim} = 475$ K and 525 K at $p \approx 10$ MPa are close to the experimental solubilities at $T = 348$ K and 398 K, respectively. Thus,

the unlike interactions are underestimated resulting from the non-polarizable model, and the same temperature scaling should be applied to compare simulation data to experimental measurements, i.e., $T_{\text{expt}} = 0.7\text{–}0.8\ T_{\text{sim}}$. I hypothesize that the χ values between linear alkane and polyol, the higher molecular weight counterparts of propane and 1,2-ethanediol, is also overestimated. Therefore, a temperature scaling factor of 0.7–0.8 is applied to all the high-χ oligomer simulation data to account for the overestimation in χ.

4.2.3 Simulation details.

MD simulations in the NpT ensemble were performed via the GROMACS 5.0.0 software [220, 221] with a time step of 2 fs. System sizes of 500–1000 molecules were used for the simulations. The temperature and pressure were controlled via the Nosé–Hoover thermostat [222, 223] with a coupling time constant of 0.4 ps and the Parrinello–Rahman barostat [224] with a 2 ps time constant, respectively. The Particle mesh Ewald method was used for the electrostatics calculation [225]. All bond lengths with hydrogen atoms were kept constant using the P-Lincs algorithm [226]. The initial structures were generated via the Monte Carlo (MC) simulations in the canonical ensemble ($T_{\text{sim}} = 3000$ K, $\rho = 0.7$ g/cm^3) for 500 MCCs using MCCCS–MN (Monte Carlo for complex chemical systems–Minnesota) software program [99]. The types of MC moves employed were similar to those used in our earlier work on oligomers, namely translation, rotation, and configurational moves [121]. In addition, the k-d tree data structure was used to accelerate the simulation [148]. After the disordered configuration was obtained, GROMACS was used to equilibrate the system at $T_{\text{sim}} = 573$ K and $p = 1$ bar. For subsequent lower temperature runs, the system was annealed from equilibrated higher temperature configuration with a rate of 2 K/ns. Each simulation was performed for at least 150 ns. The first 50 ns of the trajectory was discarded and the remainder was used for the

analysis. The uncertainties were calculated from four independent simulations with a block size of 20 ns for each run, and were reported as 95% confidence interval. VMD and Taychon were used to visualize the simulation trajectories and make the snapshot images [227, 228]. Trajectory analyses were enabled via the TRAVIS trajectory analyzer [229].

To avoid the incommensurability effect [67, 164], all of the simulations were performed for each system using an anisotropic orthorhombic simulation box, where the three dimensions of the simulation box were allowed to fluctuate independently to accommodate mesophases that are anisotropic in at least two dimensions (e.g. lamellae, perforated lamellae, and cylinders). In addition, simulations initiated from ordered mesophases were performed to assess the inaccuracy in determining the phase transition due to hysteresis. It turns out that the hysteresis is less than 30 K for the largest system size used in this study (see Figure 4.3).

Furthermore, to avoid being trapped in a local minimum, four independent simulations were performed starting from different configurations, and all of them converge to the same morphology with the same domain spacing. The excellent agreement between independent simulations indicates that the simulations reproducibly lead to the same morphology. It should be noted that the relative box lengths are not the same for all simulations; that is, the ordered morphology can be achieved through different arrangements with respect to the periodic simulation cell.

4.3 Results and Discussion

4.3.1 Characterization of the complex mesophases observed.

The morphologies of the solvent-free oligomers are examined via MD simulations. After an initial exploration of the design space, I focus on oligomers that are composed of

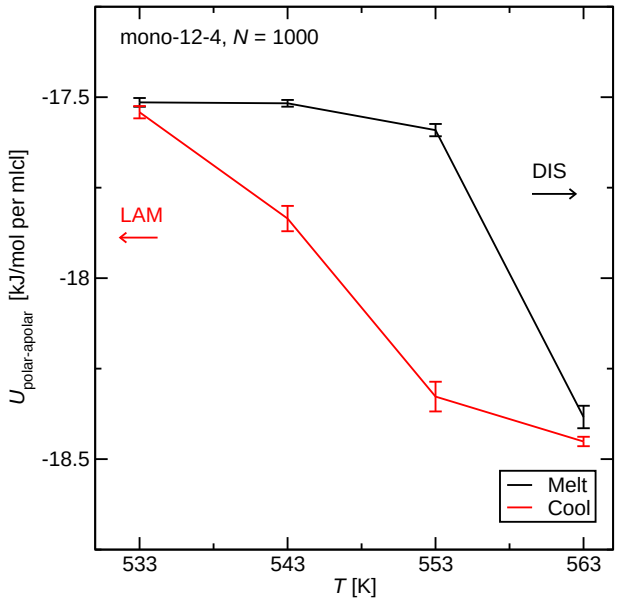

Figure 4.3: The interaction between polar and apolar blocks $U_{\mathrm{polar-apolar}}$ as a function of temperature T_{sim} for mono-12-4 oligomers. Black and red lines indicate the interaction evolution along melting and cooling, respectively.

four hydroxyl groups ($y = 4$) and up to three alkyl tails at $T_{\mathrm{sim}} = 493$, 523, and 573 K. The molecular structures and representative simulation snapshots of the mesophase at $T_{\mathrm{sim}} = 493$ K are shown in Figure 4.1. The morphologies and volume fractions of all the oligomers can be found in Table 4.1 and Table 4.2.

A variety of mesophases are observed, including lamellae, perforated lamellae, hexagonally packed cylinders, disordered micellar phases, and disordered phases. As illustrated in Figure 4.1b, one-arm (mono-) oligomers with a short hydrophobic block ($x = 12$) can phase-separate into lamellar domains. Similarly, mono-oligomers having a longer alkyl block ($x = 14$–26) can form perforated lamellae. When the length of the alkyl tail further increases to $x = 28$, the long-range ordering vanishes and the morphology transitions into the disordered phase. For two-arm (di-) oligomers, hexagonally-packed cylinders are observed in the simulations when $x = 17$–29. When the number of arms N_{arm} further

Table 4.1: Specific density ρ, volume fraction f, and morphology for mono-oligomers investigated. The abbreviations for morphologies are the same as in Figure 4.1.

Oligomer	T_{sim} [K]	ρ [g/cm^3][a]	f ($\pm$ 0.02)	Morphology
mono-12-4	573	0.8039_8	0.69	DIS
	523	0.8596_2	0.71	LAM
	493	0.8857_3	0.71	LAM
mono-14-4	573	0.7905_1	0.72	DIS
	493	0.8663_1	0.74	PL
mono-16-4	573	0.7794_2	0.75	DIS
	523	0.8273_4	0.76	PL
	493	0.8580_{10}	0.77	PL
mono-18-4	573	0.7695_1	0.77	DIS
	493	0.8483_1	0.79	PL
mono-20-4	573	0.7636_1	0.79	DIS
	523	0.8084_3	0.80	HEX-PL
	493	0.8313_6	0.80	PL
mono-22-4	573	0.7574_1	0.80	DIS
	493	0.8256_1	0.82	PL
mono-24-4	573	0.7527_1	0.82	DIS
	493	0.8200_1	0.83	PL
mono-26-4	573	0.7483_1	0.83	DIS
	493	0.8138_1	0.84	PL
mono-28-4	573	0.7435_1	0.84	DIS
	523	0.7847_1	0.85	DIS
	493	0.8075_1	0.85	DIS
mono-32-4	493	0.7991_3	0.87	DIS

[a]The statistical uncertainties in the last digit are given as the subscripts.

increases to three (tri-), disordered spherical micelles emerge from the simulations for all four chain lengths investigated (x = 13–28).

The presence of these mesophases is confirmed by the peak patterns in the structure factors shown in Figure 4.4. In addition, the domain period d can be calculated from the peak positions in the structure factor through the relationship $d = 2\pi/q^*$, where q^* is the characteristic peak position determined from the peak pattern. For lamellae, d is equal to the center-to-center spacing of the lamellae, L ($d = L = L_{\text{polar}} + L_{\text{nonpolar}}$, in which

Table 4.2: Specific density ρ, volume fraction f, and morphology for di- and tri-oligomers investigated. The abbreviations for morphologies are the same as in Figure 4.1.

Oligomer	T_{sim} [K]	ρ [g/cm^3][a]	f ($\pm$ 0.02)	Morphology
di-13-4	573	0.7875_2	0.71	DIS
	493	0.8596_2	0.73	DIS
di-17-4	573	0.7649_1	0.76	DIS
	523	0.8098_2	0.77	DIS
	493	0.8369_2	0.78	HEX-CYL
di-21-4	573	0.7512_2	0.80	DIS
	493	0.8213_2	0.81	HEX-CYL
di-29-4	573	0.7344_1	0.84	DIS
	523	0.7753_2	0.85	DIS
	493	0.8012_2	0.86	HEX-CYL
tri-13-4	493	0.8773_4	0.73	DIS-MIC
tri-16-4	573	0.7779_3	0.75	DIS-MIC
	523	0.8237_3	0.76	DIS-MIC
	493	0.8489_2	0.77	DIS-MIC
tri-22-4	493	0.8191_2	0.82	DIS-MIC
tri-28-4	573	0.7370_1	0.84	DIS-MIC
	523	0.7788_2	0.85	DIS-MIC
	493	0.8033_2	0.85	DIS-MIC

[a]The statistical uncertainties in the last digit are given as the subscripts.

L_{polar} and L_{nonpolar} are the widths of each lamella, respectively), and for hexagonally-packed cylinders, d is the 100 interplanar distance. The center-to-center spacing of cylinders L can be calculated via $d = \sqrt{3}L/2$. From the structure factor data, d for (perforated) lamellae structures at $T_{\mathrm{sim}} = 493$ K varies from 3.0 to 3.9 nm when $12 \leq x \leq 20$, and the d for hexagonally-packed cylinders ranges from 2.1 to 2.6 nm for $17 \leq x \leq 29$. The domain periods d of 3.0 nm for lamellae and 2.1 nm for cylinders are the smallest separations reported in the experimental literature for "all-organic" block polymers/oligomers [202, 22]. The corresponding domain sizes (defined as the lamellar width L_{polar} for lamellae, and the cylinder diameter D_{cyl} for hexagonally-packed cylinders) estimated from d and volume fraction f are as small as 1 nm.

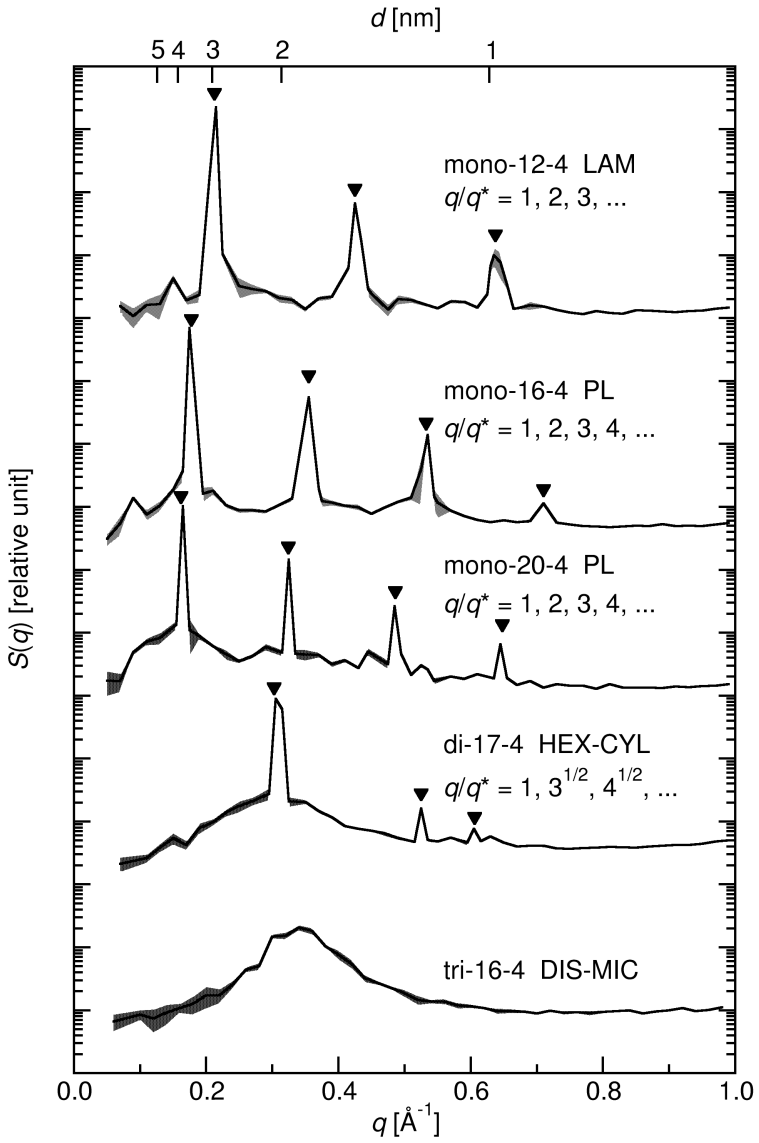

Figure 4.4: Structure factors for representative oligomers at $T_{\rm sim} = 493$ K in logarithmic scale. The data are shifted vertically for clarity. The shadow indicates uncertainties of the data. The arrow symbols indicate the peak patterns corresponding to ordered structures.

The phase behavior of these HCBOs is unique as it stands in between that of solvent-free nonionic surfactants and block polymers. On the one hand, unlike diblock polymers,

mono-oligomers can self-assemble into lamellae and perforated lamellae, but without the formation of cylindrical or spherical structures (but note that long-chain primary alcohols, such as 1-octanol, are known to form cylindrical micelles [230]). On the other hand, compared to commonly-used solvent-free nonionic surfactants, the transition temperatures from the liquid crystalline state to the isotropic liquid state of these oligomers are much higher, which implies that χ is much greater for these oligomers [209, 169, 230]. The unique aspects of block oligomer phase behavior suggests that molar mass and chain architecture both play a pivotal role; the remainder of this chapter is focused on further understanding of their phase behavior.

4.3.2 The effects of chain architecture on interfacial curvature and hydrogen bonding.

To elucidate why high-χ oligomers with multiple arms can self-assemble into cylindrical and spherical structures, I first focus on how N_{arm} shapes the phase behavior through its influence on interfacial curvature. As illustrated in the simulation snapshots in Figure 4.5a, the polar-apolar interface is more curved as N_{arm} increases, likely resulting from a bulkier alkyl tail at the interface. The shape of the interface can be quantitatively described by the angle between the end-to-end vectors of the neighboring alkyl arms (the alkyl arm is defined as one linear alkyl chain which stems from the linker carbon at the interface, e.g., each di-oligomer has two alkyl arms) as indicated in Figure 4.5b. For three sets of oligomers that have approximately the same alkyl chain length ($x = 16$ for mono- and tri-, and $x = 17$ for di-oligomers) but different values of N_{arm}, the angular distributions are clearly distinct. For mono-oligomers, a maximum at $\phi = 180^\circ$ suggests that the parallel orientation is dominant, which is in agreement with a flat polar-apolar interface. For di-oligomers, the peak height at $\phi = 180^\circ$ decreases dramatically and the angle between two neighboring arms are more randomly distributed. This packing

results in a more curved polar-apolar interface and thus the formation of cylinders, since the random orientation corresponds to neighboring arms around the cross section of the cylinder and the parallel orientation can likely be attributed to the arms along the longest axis of the cylinder. For tri-oligomers, the parallel orientation is further suppressed and the distribution becomes nearly uniform with a slightly higher bump at around $\phi = 100°$. The flat orientational distribution corresponds to a more spherical geometry as there is no preferred orientation for chain packing. Therefore, the increase in N_{arm} can effectively engender a greater curvature at the polar-apolar interface and thus a more curved morphology is favored when $N_{\text{arm}} > 1$ (cylinders for $N_{\text{arm}} = 2$ and spheres for $N_{\text{arm}} = 3$). This result is consistent with the trend predicted from the molecular packing parameter [231], as well as the mean-field prediction of miktoarm block polymer phase diagram [232, 233, 234, 200, 235].

In addition to the interfacial curvature engendered by the bulkiness of multiple arms, the second crucial factor that facilitates or frustrates the formation of ordered mesophases is the hydrogen bonding capability (equivalently, head group interactions in surfactant theory, or χ in block polymer theory). Figure 4.6a shows the number of hydrogen bonds per molecule N_{hb} at $T_{\text{sim}} = 493$ K. Stronger hydrogen bonding capability implies weaker head-tail interactions relative to head-head interactions, and equivalently, a higher effective χ parameter between two blocks. However, it is noteworthy that the hydrogen bonding is the result of complexation events, and it is essential to use more complicated models to quantitatively capture the effect of hydrogen bonding on the effective χ parameter [236]. One can see from Figure 4.6 that N_{hb} values for mono-oligomers using both of the hydrogen-bond criteria show a continuous decrease as a function of alkyl chain length x. In contrast, N_{hb} values for di-oligomers show no significant chain-length dependence. As for tri-oligomers, N_{hb} decreases slightly as x increases. This is likely due to the different morphologies of the mesophases present for mono-, di-, and

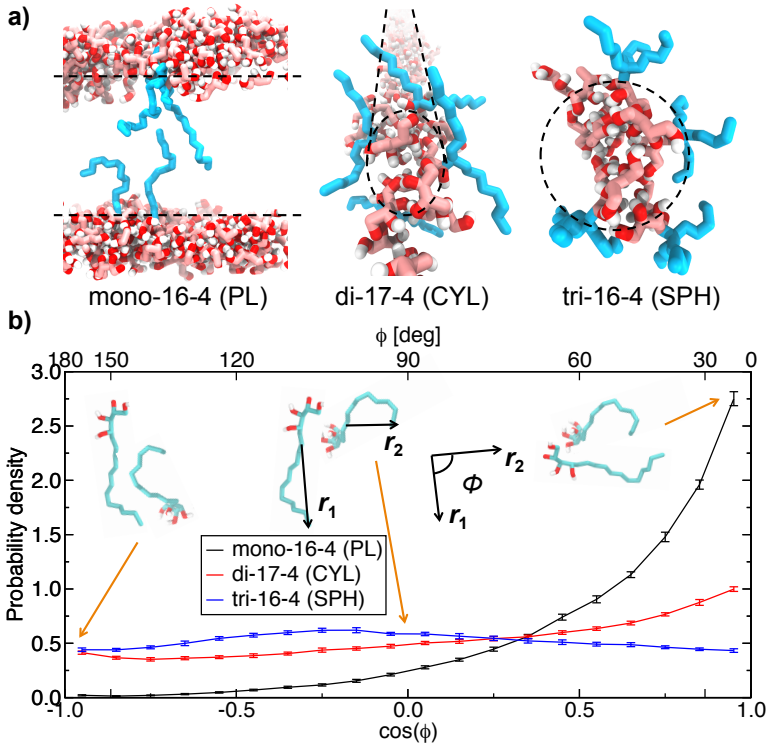

Figure 4.5: (a) Simulation snapshots of selected alkyl tails (cyan) conformations at the polar-apolar interface (black dashed lines). (b) Angular distributions of the neighboring alkyl arm end-to-end vectors for three sets of oligomers at $T_{\mathrm{sim}} = 493$ K. The schematic drawings depict the anti-parallel, perpendicular, and parallel orientation of the neighboring arms.

tri-oligomers. For mono-oligomers, as x increases, the polar lamellae contain more holes as the volume fraction of the polar block decreases (see simulation snapshots of polar lamellae in Figure 4.6). The increase in interfacial area leads to a larger fraction of the hydroxyl groups being at the interface and, hence, results in a significant decrease in N_{hb}. In contrast, di-oligomers form cylindrical structures, where the change in interfacial area as a function of x is relatively small (the diameter of the cylindrical polar region changes from 1.0 nm for di-17-4 to 1.2 nm for di-29-4). This small change in diameter results in the insensitivity of N_{hb} to the chain length. As for tri-oligomers,

they form the disordered micellar phase, which leads to a chain-length dependence in between mono- and di-oligomers. Increasing the tail length for the tri-oligomers leads to a decrease in the aggregation number and, hence, an increase in the interfacial area. In addition, the fraction of free tri-oligomers that are not involved in hydrogen-bonded aggregates increases with increasing x. Together, the two effects lead to a slight decrease in N_{hb}.

However, the effect of chain architecture on N_{hb} becomes more pronounced when N_{arm} further increases to three. N_{hb} drops by 12% for tri-16-4 oligomers compared to that of the mono-16-4 oligomers, more than a factor of two greater than the decrease from mono- to di-oligomers. The decrease is likely due to a combination of phase change and steric hindrance imposed by the three bulky arms that suppresses the presence of other hydroxyl groups near the interface. The steric hindrance works against the enthalpic driving force towards phase separation, and the resulting disordered phase further lowers N_{hb}. This hypothesis is supported by the most significant decrease in N_{hb} for the hydroxyl group closest to the alkyl tail(s) among all four hydroxyl groups, as shown in Figure 4.6b. For oligomers with two different alkyl tail lengths, the hydroxyl group closer to the alkyl tail(s) suffers more from the loss of hydrogen bonds. The first hydroxyl group in the tri-oligomers has over 20% fewer hydrogen bonds compared to that in the mono-oligomers with the same chain length, whereas the decrease in N_{hb} for the last two hydroxyl groups stays below 6%. Therefore, the reduction in hydrogen bonding for tri-oligomers results in their inability to form any ordered mesophase. In conclusion, hydrogen bonding is one of the main driving forces towards phase separation, and tuning the chain architecture is one effective approach to control hydrogen bonding and thus engineer the mesophase formation.

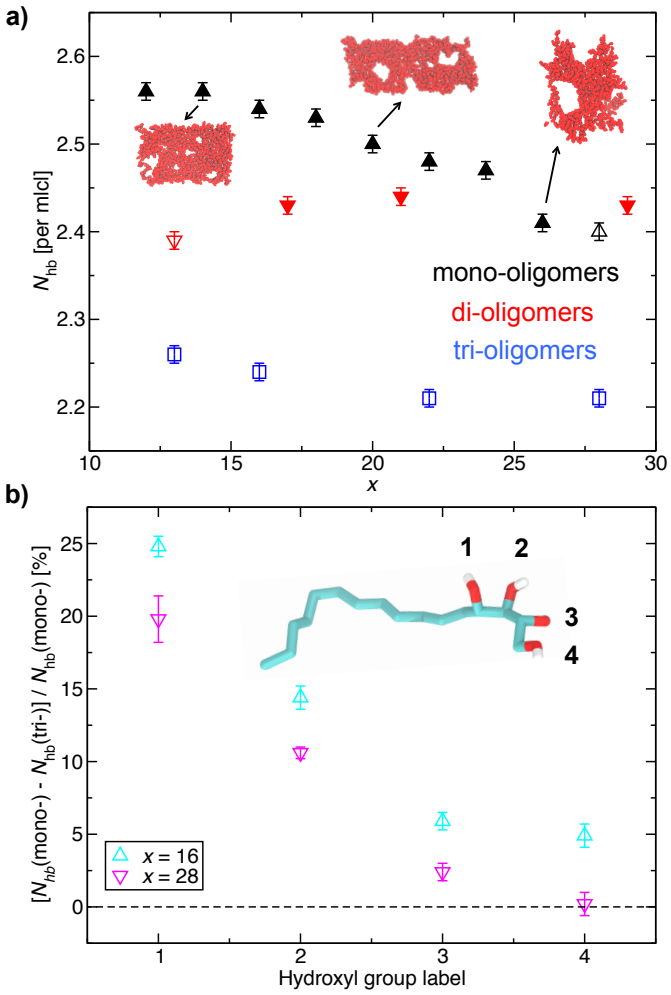

Figure 4.6: (a) The number of hydrogen bonds per molecule $N_{\rm hb}$ for three types of oligomers at $T_{\rm sim}$ = 493 K. The filled and open symbols correspond to oligomers in the ordered and disordered phases, respectively. Snapshots of perforated lamellae for mono-14-4, mono-20-4, and mono-26-4 are shown. (b) The relative decrease in $N_{\rm hb}$ from mono-oligomers to tri-oligomers as a function of hydroxyl group position for oligomers with two different alkyl tail lengths at $T_{\rm sim}$ = 493 K.

4.3.3 The effect of chain length on molecular conformation and packing.

The key to interpreting the absence of cylindrical and spherical mesophases for mono-oligomers is the understanding of how the alkyl chain length alters the molecular conformation and thus the stability of mesophases. Figure 4.7a illustrates the scaling behavior of the end-to-end distance of the alkyl arm as a function of its chain length $R_{ee} \propto x^{\nu}$, where ν is the Flory exponent. Note that the end-to-end carbon distance of the polar block is always about 3.3 Å (a predominantly *gauche* conformation), and thus, it plays a less important role in affecting the bulk morphology. In the disordered state at $T_{sim} = 573$ K, ν is found to decrease as x increases. It ranges from 0.8 for shorter chains to 0.6 for longer ones, which is greater than the scaling exponent of 0.5 for a long polymer in its pure melt. This indicates that the oligomers are more rigid compared to their polymer counterpart, which aligns well with results from our earlier studies on olefin oligomers [121]. At $T_{sim} = 493$ K, the Flory exponent decreases continuously as a function of x, from 0.9 to 0.6, which reveals that particularly the shorter chains become more rigid at lower temperatures. It is noteworthy that the Flory exponents of the short oligomers ($x \leq 20$) are greater than those predicted by the mean-field theory for block polymers in the ordered state ($\nu = 0.5$ from the weak segregation theory [237, 238] and $\nu = 0.67$ from the strong segregation theory [238]).

The greater persistence of the chains implies a transition from flexible coils to rigid rods as the alkyl tail chain length decreases. For flexible block polymers, perforated lamellae are considered thermodynamically unstable [239, 240]. However, the presence of perforated lamellae has been reported by both theoretical [241] and experimental [242] studies for rod-coil block polymers, which is consistent with the mono-oligomer phase behavior. The enhanced stability of perforated lamellae is schematically explained in Figure 4.7b. When the chains are sufficiently long, the scaling behavior ($\nu \approx 0.6$)

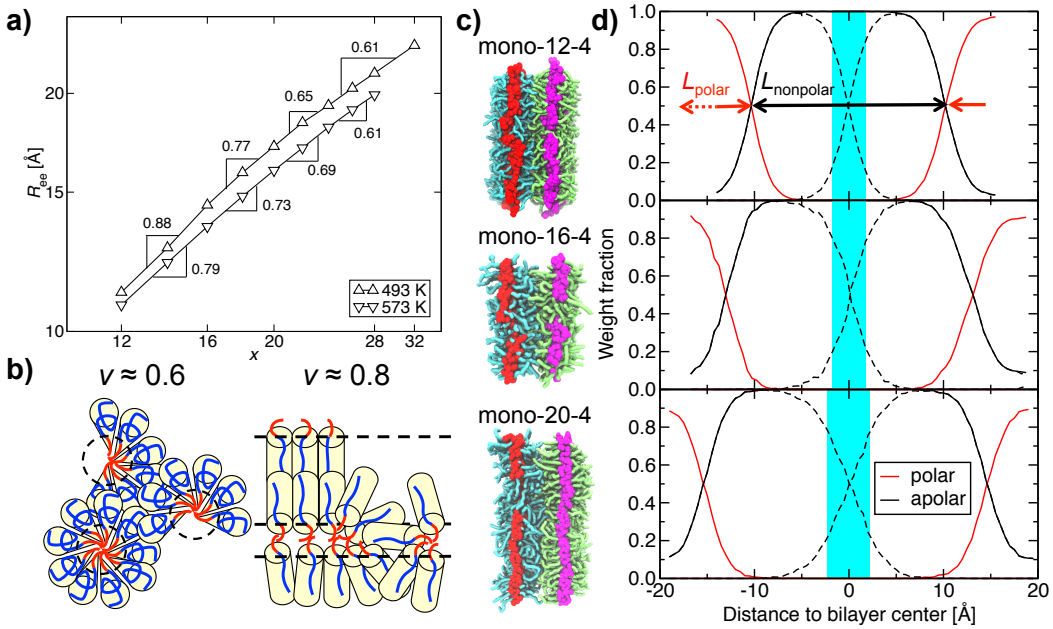

Figure 4.7: (a) End-to-end chain distance of the alkyl arm (R_{ee}) as a function of x for mono-oligomers, on logarithmic scale. The slopes for the two or three neighboring data points, which correspond to the Flory exponent ν, are given. (b) Schematic drawing of how the chain conformation affects packing, where red and blue lines correspond to polar and apolar blocks, respectively, and dashed lines indicate the interface. (c) Simulation snapshots of representative (perforated) lamellae for three oligomers at $T_{sim} = 493$ K. Atoms in the polar block are rendered in red and magenta for two leaflets (chains that have their head groups in the same cluster form one leaflet), respectively, and apolar groups that stem from two leaflets are rendered in cyan and green, respectively. (d) Composition profile along the axis perpendicular to the lamellae plane corresponding to the three oligomers shown in (c). The composition profiles for apolar blocks are shown separated for each leaflet (dashed lines) and for both leaflets combined (solid lines). The cyan-colored region indicates the interdigitation region where apolar weight fractions of both leaflets are above 0.2.

approaches that of the polymer melt ($\nu \approx 0.5$). In this case, the chains can be modeled as cones that can effectively pack into hexagonal structures. In contrast, when the oligomer chains are more rigid ($\nu \approx 0.8$), the rod-like chains lack the ability to form hexagonally-packed structures and instead prefers layered structures [241]. At this range of chain length, the oligomer phase behavior resembles that of a liquid crystal, where cylindrical and spherical structures can result in packing defects and thus are energetically not favored [243].

The effect of chain conformation on the layered structure packing is further confirmed by the simulation snapshots of the bilayer structure and its composition profile depicted in Figure 4.7c and Figure 4.7d, respectively. From the simulation snapshots, alkyl chains tend to stretch out from the interface to the center of bilayer without much interdigitation between the chains around the center of bilayer. This observation is supported by the composition profile along the axis perpendicular to the lamellae plane. As x increases from 12 to 20, the thickness of the alkyl layer L_{nonpolar} grows from 2.1 to 3.0 nm. However, the change in the thickness of the alkyl interdigitation region (shown in cyan in the figure) is less pronounced, from 0.35 nm for $x = 12$ to 0.45 nm for $x = 20$. This implies that the chains attempt to avoid interdigitation, which would potentially perturb their extended conformation and parallel alignment. This observation confirms the rod-like chain conformation and its effect on the phase behavior.

It is intriguing whether the patterning of such small domains will result in larger line edge roughness (LER), or even defects in the lamellae plane due to the Plateau-Rayleigh instability [244]. LER (3σ) is computed for the bulk self-assembled mono-12-4 oligomer at $T_{\text{sim}} = 493$ K and a value of 2.0 ± 0.3 Å is obtained. The LER is around 20% of the domain width, which is sufficient for most nanolithography applications [245, 246, 247].

Figure 4.8a illustrates the phase map that includes all the oligomers investigated in this study. The mean-field χ parameter, χ_{MF}, is obtained from fitting our simulation

results to the order-disorder phase boundary from self-consistent field theory [248, 240, 249, 232, 233], which results in the following:

$$\chi_{\mathrm{MF}}(T_{\mathrm{sim}}) = \frac{2150}{T_{\mathrm{sim}}} - 1.1 \tag{4.1}$$

for mono-oligomers, and

$$\chi_{\mathrm{MF}}(T_{\mathrm{sim}}) = \frac{3440}{T_{\mathrm{sim}}} - 2.8 \tag{4.2}$$

for di-oligomers. For both cases, a typical reference volume of 118 $\mathrm{\AA}^3$ is used. There is insufficient data to fit χ_{MF} for tri-oligomers since all the tri-oligomers covered in this study are in the disordered phase.

The comparison of χ and the achieved domain period d for the mean-field boundary (assuming a typical statistical segment length of $b = 0.7$ nm for convenience of comparison), a few selected experimentally available block polymers [182, 193, 194, 195, 196, 197, 190, 198, 199], and the oligomers from this work are shown in Figure 4.8b. χ_{MF} can be as large as 4.3 and 5.8 at $T_{\mathrm{sim}} = 400$ K ($T_{\mathrm{expt}} \approx 300$ K) for mono-oligomers and di-oligomers, respectively. These oligomers have χ_{MF} values that are one to two orders of magnitude greater than those of other "all-organic" block oligomers. For example, the χ_{MF} of poly(styrene-*b*-methyl methacrylate), one of the most widely studied diblock polymer for self-assembly, is around 0.03 at $T = 300$ K [250], while the largest χ_{MF} value for "all-organic" diblock polymers reported experimentally is less than 0.5 at $T = 300$ K [196, 193]. The greater χ for these oligomers allows the decrease in N and thus enables the formation of 1-nm domains.

One should note, however, that use of χ_{MF} is a convenient but flawed approach. Nevertheless, presenting some of the data and discussion in terms of the χ parameter allows for comparison of our results to those from the standard diblock copolymer model applicable to blocks consisting of at least a few Kuhn segments and not involving strong

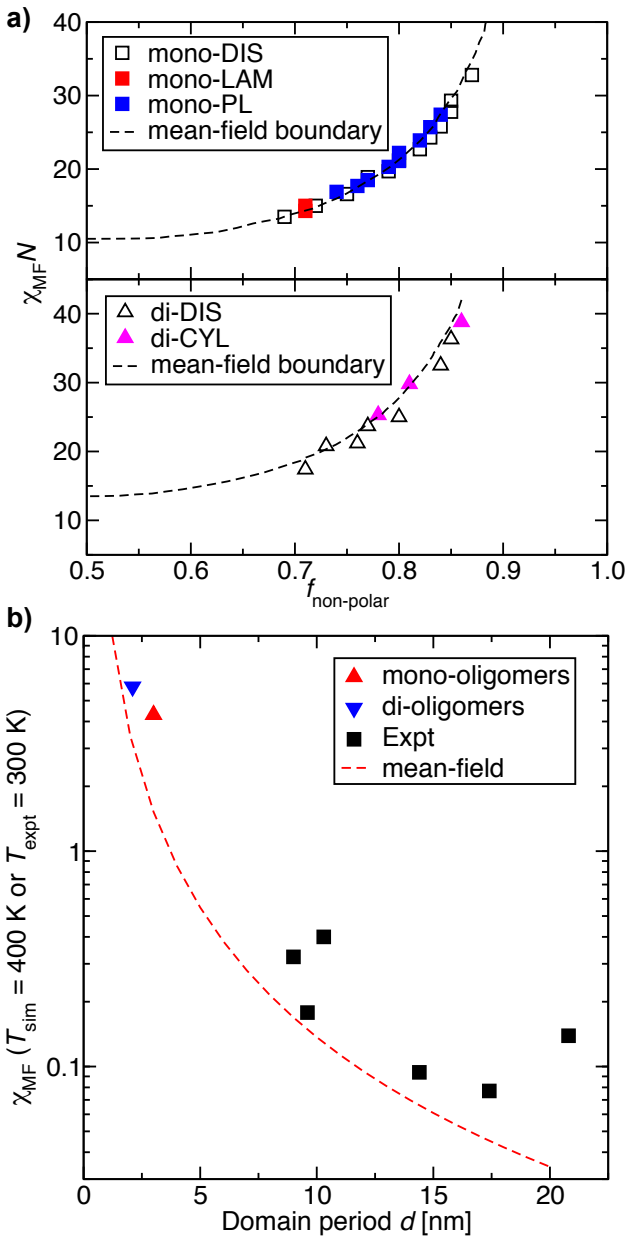

Figure 4.8: (a) The phase map for the oligomers investigated shown in symbols. χ_{MF}, the mean-field χ, is fitted to the order-disorder phase boundary from the self-consistent mean-field theory [248, 240, 249, 232, 233] represented by the dashed lines. (b) The χ_{MF} as a function of domain period d for mean-field boundary ($\chi_{\mathrm{MF}} > 13.7/(d/\mathrm{nm})^2$, the red dashed line), available "all-organic" materials (black symbols) [182, 193, 194, 195, 196, 197, 190, 198, 199], and oligomers from this work (triangles). χ_{MF} values are from extrapolation to $T_{\mathrm{expt}} = 300$ K or $T_{\mathrm{sim}} = 400$ K, and the simulation domain period values are from the $T_{\mathrm{sim}} = 493$ K data.

directional interactions. There are two major factors that are not included in the fitting of χ_{MF}. First, complexation interactions such as hydrogen bonding are not considered in the Flory–Huggins theory formulation [236]. Consequently, χ_{MF} is a function of y due to the cooperative nature of hydrogen bonding. For example, mono-6-2 and mono-9-3 oligomers were simulated at $T_{\mathrm{sim}} = 493$ K. Both are found to form homogeneous disordered phases, which leads to the conclusion that $\chi_{\mathrm{MF}}(y = 4) > \chi_{\mathrm{MF}}(y = 3)$. Second, the effects of short chain length and semiflexibility are not taken into account in the fitting of mono- and di-oligomers data using the mean-field AB diblock copolymer model and the mean-field AB_2 star polymer model, respectively [251, 69]. A recent computational work by Mao et al. [69] has discussed the effects of finite chain length and semiflexibility on the order-disorder transition point χN_{ODT} for symmetric diblock polymers. According to their results [69], χN_{ODT} differs by around 20% from the mean-field theory for symmetric diblock copolymers that have similar chain length and semiflexibility as HCBOs. Despite that the quantitative determination of χ is hampered by the lack of $\chi N_{\mathrm{ODT}}(f)$ theory data for oligomeric semiflexible chains, these valuable data demonstrate that the χ_{MF} values from the semi-quantitative fitting in our manuscript are only for comparison purposes and not accurate reflections of only the interactions between the two blocks in standard block polymer models.

Despite the loss of predictive power from the mean-field theory, the deviations from the Flory–Huggins theory framework can be more of a blessing than a curse. Novel applications of HCBOs can arise from their potential to access domain morphologies that are beyond the reach of flexible diblock polymers. For example, researchers have observed a variety of lamellae phases including perforated lamellae, wavy lamellae, zigzag lamellae, and arrowhead lamellae in the rod-coil block polymer phase map [252, 241]. Therefore, further simulation, experimental, and theoretical studies are needed to elucidate the phase behavior and additional applications of these HCBOs.

4.4 Conclusions

In this study, a series of high-χ block oligomers were designed with various chain lengths, volume fractions, and chain architectures. From MD simulations, these oligomers can self-assemble into a variety of mesophases including lamellae, perforated lamellae, and hexagonally-packed cylinders. The smallest domain periods in lamellae and cylinders are 3.0 nm and 2.1 nm, respectively, and the smallest polar domain sizes are around 1 nm. These are the smallest reported in the literature so far for "all-organic" block oligomers. At this length scale, the phase behavior stands in between that of block polymer and that of solvent-free surfactant, which arises from the interplay between both "block-polymer factors" (χ, N, and f) and "surfactant factors" (head group interactions, chain flexibility, and interfacial curvature). The chain architecture can greatly affect the interfacial curvature and hydrogen bonding, which consequently facilitates or frustrates the phase separation and the resulting mesophases. The chain conformation, on the other hand, is more stretched than that of its polymer counterpart, which is believed to be the cause for the unusually high stability of perforated lamellae phase. These insights provide a route to engineer similar HCBOs at the sub-5 nm length scale to the desired morphology and domain separation for applications such as nanolithography.

Chapter 5

Using the k-d Tree Data Structure to Accelerate Monte Carlo Simulations

5.1 Introduction

Molecular simulation techniques, such as molecular dynamics (MD) and Monte Carlo (MC), are widely used to study structural, thermodynamic, and transport properties of chemical systems [253, 219, 44]. The efficiency of these simulation methods is pivotal in accessing larger length and times scales and more complex force fields in order to provide more precise and accurate predictions. The computation of the non-bonded interactions is the most time-consuming part in particle-based N-body simulations [219, 44, 254, 255]. In many molecular simulations, non-bonded interactions are pairwise-additive and truncated at a spherical cutoff distance, r_{cut}. In this case, an efficient algorithm to

This chapter is reproduced in part with permission from (Q. P. Chen, B. Xue and J. I. Siepmann, "Using the k-d Tree Data Structure to Accelerate Monte Carlo Simulations", *J. Chem. Theory Comput.* **2017**, *13*, 1556-1565, American Chemical Society)

perform a range search (defined as searching for interaction sites that are within r_{cut} from one specific interaction site) becomes pivotal to the performance of the simulation program.

The standard (complete loop) implementation [219] requires one distance calculation between the specific interaction site and every other interaction site in the simulation box; therefore, it scales as $\mathcal{O}(N)$, where N is the number of molecules in the system (for multi-component mixtures, the number of interaction sites, N_{site}, determines the scaling). In MD simulations, the range search is often accelerated by using the Verlet neighbor list and its embellishments [256, 257, 258, 259, 260, 261, 262, 263, 264, 254, 265]. Nevertheless, these neighbor lists are not suitable for many MC simulations because of more frequent neighbor list updates resulting from advanced MC moves where particles can be moved over large distances or transferred to another simulation box.

Some MC simulation programs utilize a center-of-mass (COM) cutoff (or a leader atom for molecular fragments) [266, 107] to enhance the efficiency of the range search. During such a range search, molecule pairs are first screened by their COM distance, and the site–site distances are computed only if the following condition is satisfied:

$$r_{\text{COM}} \;\leq\; r_{\text{cut}} + \Delta r_{\text{intra},i} + \Delta r_{\text{intra},j} \tag{5.1}$$

where r_{COM} is the distance between the COMs of molecules i and j; $\Delta r_{\text{intra},k}$ is the maximum possible distance between the COM and any interaction site in molecule k. This approach reduces $N_{\text{site},i}N_{\text{site},j}$ distance calculations to only one for molecule pairs that do not satisfy the criterion given by eq 5.1, where $N_{\text{site},k}$ is the number of interaction site of molecule k. For molecules with three or more interaction sites, the increase in the computational cost associated with the calculation of r_{COM} is only marginal for a pair that satisfies eq 5.1. Many MC simulation programs keep track of the COM positions anyway

because these are used as centers for molecular rotations and to maintain molecules in the central simulation box (thereby allowing for faster implementation of the periodic boundary condition). Due to the increase of $\Delta r_{\text{intra},k}$, the COM cutoff approach works best for molecules with multiple beads but relatively compact architecture.

Another approach for speeding up the range search is the use of linked-cell lists, where the simulation box is spatially divided into equally spaced grid cells [219, 267, 268, 269]. If the size of the grid cell, L_{cell}, satisfies the following condition:

$$r_{\text{cut}} \;\leq\; L_{\text{cell}} \;\leq\; L_{\text{box}} \,/\, i_{\geq 4} \tag{5.2}$$

where L_{box} is the length of the simulation box and $i_{\geq 4}$ is an integer greater than or equal to 4, then the range search can be accelerated by searching only 27 cells (the cell that the interaction site is residing in, and its 26 neighboring cells) rather than all the $(i_{\geq 4})^3$ cells for a cubic simulation box. Despite the additional computational and memory costs associated with book-keeping in which cell every molecule resides, a range search using the linked-cell list scales as $\mathcal{O}(1)$, if r_{cut} is independent of N. In addition to a cell size that is equal or slightly larger than r_{cut}, a much smaller cell size allowing at most one bead in each cell can also be utilized to reduce the number of distance calculations in the range search [264, 270, 271, 272, 273, 274, 267]. However, the use of smaller cell sizes often leads to intense memory usage and thus is not considered in this study.

In this work, the use of the k-d tree data structure is proposed as a novel approach to accelerate the range search in MC simulations. The standard array implementation [219] to store particle coordinates supports coordinate update and range search scaling as $\mathcal{O}(1)$ and $\mathcal{O}(N)$, respectively. In contrast, the k-d tree is a binary tree data structure that supports these two operations both in $\mathcal{O}(\log_2 N)$ time (assuming r_{cut} is independent of N) [275]. When the system size is sufficiently large, the speed-up from the faster

range search can outweigh the additional cost associated with updating coordinates in the tree. In the literature, the k-d tree is mostly used for clustering interaction sites in astronomical simulations or as a fast multipole method in computing long-range interactions [276, 277]. In this study, the k-d tree data structure is implemented in a MC simulation software package, where its efficiency enhancement is expected to come solely from a more efficient range search algorithm. The applicability of the k-d tree is studied via efficiency tests on a wide range of systems and ensembles.

The chapter is organized as follows: In the next section, a brief introduction of the k-d tree data structure is given. The third section focuses on the performance tests comparing the k-d tree to complete loop, COM cutoff, and linked-cell list approaches. These efficiency tests provide rules of thumb on whether to use the k-d tree data structure for a type of molecular system. They also help us to gain a better understanding of the k-d tree performance and facilitate algorithm and data structure design in the future.

5.2 Theory and Computational Details

5.2.1 Background on the k-d tree data structure.

The k-d tree is a binary tree data structure that supports insertion, deletion, and range search in $\mathcal{O}(\log_2 N)$ time [275]. In a k-d tree, one node contains information for one interaction site (e.g., coordinates, type of site, and partial charge) in the interacting system, and it has at most two child nodes. Each tree level is associated with a cutting dimension which is alternatively x, y, or z for a system with Cartesian coordinates. For every node, the coordinate in the cutting dimension is always greater than that of any node in its left sub-tree, and smaller than that of any node in its right sub-tree. Geometrically, this means that every node poses a binary division of the simulation cell in x, y, or z dimension. In other words, each sub-tree is bounded by a rectangular

cuboid, and the size of this bounding cuboid decreases as the tree is traversed from the root to the leaves.

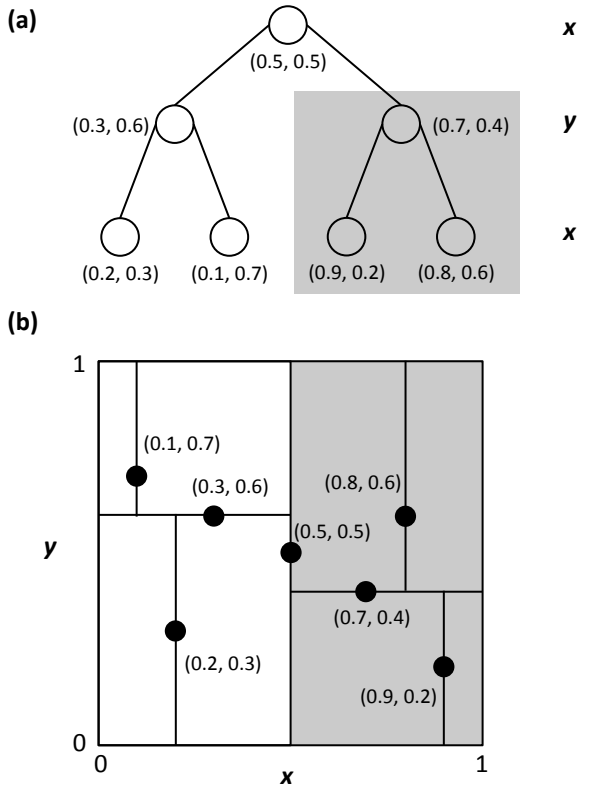

Figure 5.1: Schematic illustration of a two-dimensional k-d tree which represents seven single-bead particles in a square simulation box without the use of the periodic boundary condition. (a) The constructed balanced k-d tree (see the next section for tree construction details), with each node representing one particle; the cutting dimension (x or y) is shown on the right side of each level. (b) The corresponding simulation box with seven particles, with vertical and horizontal lines representing geometric sub-divisions in the cutting dimensions and with cells (bounding boxes) representing the left/right or up/down sub-trees of a certain node. For example, the shadowed cell represents the right sub-tree of the root node, also shown in gray background in (a).

A two-dimensional k-d tree together with a toy example system is shown in Figure 5.1 and illustrates how molecular simulations can benefit from the use of the k-d tree.

In this example, the simulation box is geometrically divided into eight smaller cells, and each cell is the bounding box of one leaf in the tree. For example, in Figure 5.1, the shadowed cell in the simulation box is the bounding box of the shadowed sub-tree. The range search in the k-d tree can be greatly accelerated by using the information of the bounding box of a sub-tree. When the range search is performed for a given node, the search in an entire sub-tree can be skipped if the shortest distance between the node and the bounding box of this sub-tree is greater than r_{cut}. In the toy example, if we are searching for all the nodes within a distance of 0.2 units away from the node (0.8, 0.6), the entire left sub-tree of the root does not need to be traversed because the shortest distance between the node (0.8, 0.6) and the bounding box of that sub-tree is 0.3 units. The reduction in the number of distance calculations can be a major source of efficiency improvement that leads to a better asymptotic scaling of the range search in the k-d tree ($\mathcal{O}(\log_2 N)$) compared to that of the array implementation ($\mathcal{O}(N)$).

5.2.2 k-d tree implementation details.

The k-d tree data structure is implemented in the in-house MC simulation software package MCCCS-MN (Monte Carlo for Complex Chemical Systems-Minnesota) [99]. There are two details worth noting for our k-d tree implementation: First, the periodic boundary condition is treated in a manner that periodic images that might interact with beads in the central box are also stored and updated in the k-d tree. Second, the parallelization of the k-d tree has not been considered in this study because the performance would depend more on the architecture of the simulation code rather than the k-d tree data structure itself. Nevertheless, it is noteworthy that the k-d tree data structure has been parallelized for both CPU [278, 279, 280] and GPU [281] architectures.

k-d tree construction. The k-d tree data structure needs to be initialized at the start of the simulation when an initial configuration is either read from the file or generated by

the software. In addition, the k-d tree needs to be reconstructed to ensure the balance of the tree and, hence, efficient insertion, deletion, and range search. During the tree initialization, the memory is allocated for every node, and then parent-child points for each node is assigned. For the tree reconstruction, the memory does not need to be reallocated because the "in-place" sorting algorithm is used (which will be discussed in the next paragraph). The k-d tree can be perfectly balanced after the initialization or reconstruction by always inserting the node whose cutting dimension coordinate is the median among all the other nodes first, which can be decomposed into the following steps:

1. Split all the nodes that have not yet been inserted into two halves based on the cutting
   dimension coordinate.
2. Put the median node into the insertion queue.
3. Recursively carry out step (1) to the smaller and greater half.
4. Insert all the nodes in the insertion queue to the tree.

The split is performed via the "in-place" `quickselect` algorithm and it scales as $\mathcal{O}(N)$ [282], resulting in an overall scaling of $\mathcal{O}(N\log_2 N)$ for the tree construction. The treatment of the periodic boundary condition (see below) might introduce beads that are periodic images and, hence, there can be a tie when it comes to sorting or comparison in the k-d tree. In this case, the coordinate of the next cutting dimension can serve as a tie breaker (e.g. comparing y coordinates if x coordinates are equal) to ensure the balance of the tree.

The treatment of the periodic boundary condition can make a significant impact on the performance. To reduce the number of nodes stored in the k-d tree, all the

periodic images are stored in the k-d tree only if their coordinates in the i^{th} dimension, r_i $(i = x, y, z)$, satisfies the following condition:

$$-r_{\text{cut}} - 2\Delta r_{\text{intra}} \leq r_i \leq r_{\text{cut}} + 2\Delta r_{\text{intra}} \tag{5.3}$$

The method is used when particles are "wrapped" into the central box on a COM basis, i.e., the coordinate of a particular bead may exceed the box length as long as the COM of the molecule does not. For other "wrapping" options, similar approaches can be adopted accordingly. The current approach ensures that only periodic images that may interact with nodes in the central box are stored in the k-d tree. Because of the storage of coordinates for additional periodic images, the faster range search of this extended k-d tree comes with the price of higher memory usage. Nevertheless, this does not affect the performance of the MCCCS–MN program for the system sizes considered here.

Node insertion and update. A new node is inserted into the tree by first traversing the tree through a series of comparisons and then being added as a new leaf in the appropriate position. The coordinate update during the course of a simulation is achieved by first deleting the old node and then inserting the new one. The deletion involves finding the node that has the minimum coordinate in the same cutting dimension in the right sub-tree (or equivalently, the maximum node in the left sub-tree), exchanging the to-be-deleted node with this minimum node (or the maximum node), and then iterating the previous two steps recursively, until the to-be-deleted node has no child node and therefore can be simply removed. The node update can result in the imbalance of the tree. For local MC moves such as translation or rotation, local updates may be designed to maintain the balance of the tree, but it is difficult to design such operations for non-local MC moves such as the particle transfer move or the aggregation-volume bias MC move [283]. Therefore, the tree reconstruction is performed to maintain the balance of

the tree after the node update, which will be discussed later.

Range search. The range search for particles within r_{cut} from any arbitrary position in the simulation box (or reference position) can be performed via the following steps:

1. Examine the current node, and include it in the "found neighbor" list if it is within r_{cut} from the reference position.

2. Compute the distance from the reference position to the left and right sub-trees of the current node. The distance from a node to a sub-tree is defined as the shortest distance between the reference coordinate and the bounding cube of this sub-tree, and it is zero when the reference coordinate is within the bounding cube. This is the closest possible distance between the reference position and every particle in the sub-tree.

3. Recursively examine the sub-tree that is "closer" in distance from the reference position starting from step (1).

4. After all of the "closer" trees are examined, recursively examine the "farther" sub-tree starting from step (1) when the distance from the reference coordinate to the sub-tree is smaller or equal to r_{cut}.

Timing scaling and tree reconstruction When the k-d tree is perfectly balanced, then the height of the tree $H_{\min} = \lceil \log_2 N \rceil$, and insertions and deletions of a single particle scale as $\mathcal{O}(\log_2 N)$, where N is the number of interaction sites including the ones in the periodic images that might interact with the interaction sites in the original simulation box. Thus, k-d tree insertions and deletions are more expensive than the coordinate update in the standard array implementation ($\mathcal{O}(1)$). However, the range search using the k-d tree scales asymptotically as $\mathcal{O}(\log_2 N)$, that is significantly improved over the $\mathcal{O}(N)$ scaling for the array implementation.

It depends on the type of MC moves whether to reconstruct the tree during the simulation: If the volume move is not used, the tree is reconstructed once the tree height exceeds a threshold value. On the other hand, if the volume move is used, the tree is not reconstructed.

Figure 5.2 indicates the evolution of tree height as a function of number of MCSs for Lennard-Jonesium simulations in the NVT ensemble at three different system sizes. The tree height grows exponentially and reaches a plateau at around twice the balanced tree height $H_{\min}$ ($H_{\min}$ = 15, 15, and 16 for N = 1500, 3000, and 9000, respectively). The effect of an imbalanced tree on the efficiency performance is described in Figure 5.3, in which the CPU time is plotted as a function of the reconstruction threshold. The k-d tree is reconstructed if $H/H_{\min}$, the ratio between the tree height and the balanced tree height, exceeds a threshold value. The CPU time first decreases as the reconstruction threshold becomes greater, and then it increases when the threshold value is greater than 1.25. The two competing factors are the additional CPU time spent on reconstructing the tree and the CPU time saved to perform tree operations on a more balanced tree. The optimal CPU time at the threshold value of 1.25 is about 10% smaller compared to that if the tree is never reconstructed. Therefore, for all of the simulations without the volume move, the tree reconstruction threshold is set to be 1.25.

From Figure 5.3, it can also be deduced that a threshold value of 1.25 leads to a tree reconstruction frequency of every few MCCs, depending on the system size (3.2 MC cycles for N = 1500, 1 MC cycle for N = 3000, and 4 MC cycles for N = 9000). Since on average there is more than one accepted volume move per MC cycle, there is no need to reconstruct the k-d tree for simulations with volume moves.

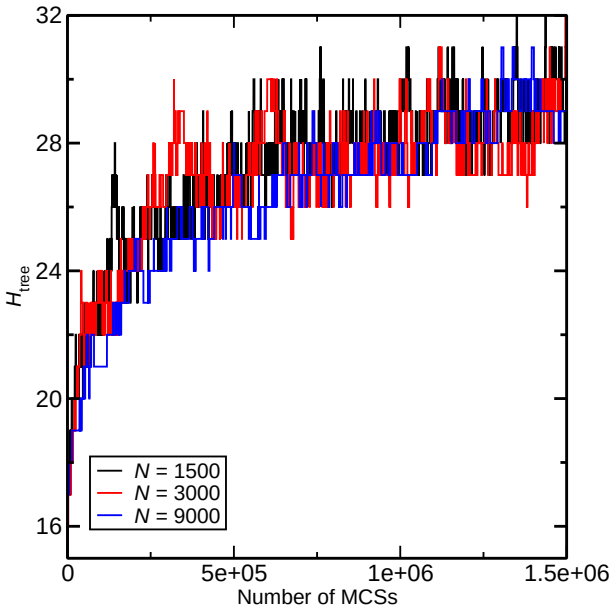

Figure 5.2: The height of the tree as a function of number of MCSs for Lennard-Jonesium simulation in the NVT ensemble ($T^* = 0.72$, $\rho^* = 0.75$, and $r^*_{\text{cut}} = 4$)

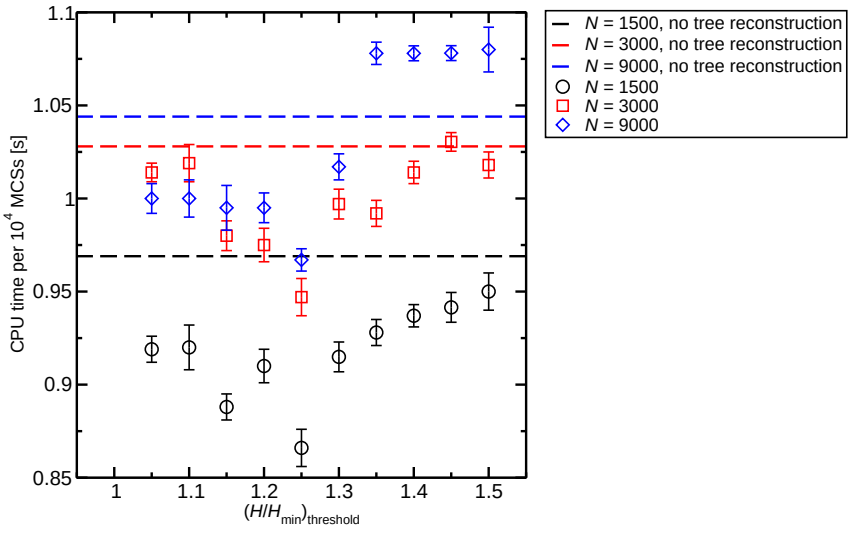

Figure 5.3: CPU timings for simulations of Lennard-Jonesium simulations in the NVT ensemble ($T^* = 0.72$, $\rho^* = 0.75$, and $r^*_{\text{cut}} = 4$) as a function of the tree reconstruction threshold $(H/H_{\text{min}})_{\text{threshold}}$. The tree is reconstructed if the reduced height of the tree H/H_{min} exceeds this value.

5.2.3 Computational details.

All the benchmark tests were performed using MCCCS-MN. Other approaches for the range search including complete loop, linked-cell list, and COM cutoff were previously implemented in MCCCS–MN, and are used for comparison with the k-d tree. Among these techniques, the linked-cell list implementation is MCCCS–MN specific with regard to the following two aspects. First, the linked-cell list is constructed on a COM basis (i.e., the whole molecule is stored in a cell based on its COM coordinate), not on an interaction site basis (the reason is that this improves efficiency for calculations of intramolecular interactions and for updating the list). The cubic cell length is determined by the following equation: $L_{\text{cell}} = r_{\text{cut}} + 2\Delta r_{\text{intra}}$. This COM-based linked-cell list allows for the concurrent use of the COM cutoff and, in theory, the performance should not differ from that of the site-based linked-cell list for molecular systems where $\Delta r_{\text{intra}} \ll r_{\text{cut}}$ (i.e., butane, ethanol, and water, but not for the high-molecular-weight alkane considered in this study). Second, the linked-cell list can only be applied to one simulation box because, for simulations of vapor–liquid equilibria, the liquid-phase box usually contains significantly more molecules and the range search in this box is the largest cost. Nevertheless, the k-d tree will be shown to be more efficient for a single simulation box and, hence, there is no need to implement the linked-cell list for multiple simulation boxes. Despite that the features mentioned here do not allow for a direct comparisons between k-d tree and linked-cell list performances in a few cases, the conclusions in the result section still hold with adequate fair comparisons between the two.

The compounds, ensembles, and system sizes considered here are summarized in Table 5.1. The compounds were selected to cover single-bead particles (Lennard-Jonesium), models with four interaction sites (n-butane, ethanol, and water), and 2,4,6,8,10,12,14,16,-18,20,22-undecamethylpentacosane as an oligomeric compound with 36 interaction sites, models without partial charges (Lennard-Jonesium and the two alkanes) and models with

Table 5.1: Systems considered in this study.

Molecule	Force field	Ensemble	N
Lennard-Jonesium	N/A	NVT	400 – 25000
Lennard-Jonesium	N/A	NVT-Gibbs	1500 – 15000
n-Butane	TraPPE–UA [95]	NpT	1500 – 6000
Ethanol	TraPPE–UA [216]	NpT	1500 – 6000
Water	TIP4P [266]	NpT	2000 – 8000
Water	TIP4P [266]	NVT-Gibbs	2000 – 8000
$C_{36}H_{72}$	TraPPE–UA [95, 96]	NpT	100 – 400

three partial charges (ethanol and water), and rigid (water) and semiflexible molecules (alkanes and ethanol). For Lennard-Jonesium, $r_{\text{cut}} = 4\sigma$ is used, whereas the standard $r_{\text{cut}} = 14$ Å of the TraPPE force field [95, 96, 216] is used for all molecular systems. The choice of r_{cut} is an important parameter that significantly affects the absolute CPU time required for a simulation, but its influence on the relative timings for different range search algorithms is limited. The types of MC moves used in these simulations include translations of molecules, rotations of molecules around the COM, volume moves scaling the COM positions, dual-cutoff coupled-decoupled configurational-bias Monte Carlo moves (here abbreviated as CBMC) [96], and CBMC particle transfer (swap) moves [55]. For the CBMC moves, a 5 Å inner cutoff is used and Coulomb interactions are excluded when Rosenbluth weights are computed, and then the weights of only the trial and old configurations are corrected to include interactions out to r_{cut} and long-range interactions in the acceptance rule [107]. In CBMC conformational and swap moves, 16 trial positions (the exception is water in the Gibbs ensemble where 32 trial positions are used for the initial insertion of the oxygen site) are generated and one of them is chosen based on the Boltzmann weight of the set of candidate positions. As is the standard practice, only the energy change caused by the displacement of a single molecule (requiring calculations for the molecule in its new and old configuration) is computed for translational,

rotational, and CBMC moves. Here, the energy of the molecule in the new configuration is computed first, and, if it results in an extremely high energy, or results in an overlap ($r_{\text{overlap}} = 0.3\sigma, 1.0$ Å and 1.2 Å for Lennard-Jonesium, alkanes, and hydrogen-bonding molecules, respectively), then the trial move is rejected without computing the energy of the molecule in the old configuration. A similar approach is used for the Rosenbluth weight and can lead to early termination of a CBMC move (the termination rate increases as more beads are attempted to be regrown in a CBMC move). For example, the early termination rates are 1.3% and 0.15% for regrowing the entire n-butane ($T = 255$ K) and ethanol ($T = 323$ K) molecules, respectively, but it increases to 83% for the entire 36-carbon oligomer propylene molecule. For the volume move, only the total energy of the trial configuration needs to be computed because that of the old configuration is always known. Analytical tail corrections for the Lennard-Jones interactions are used for all systems [219, 44]. The Ewald summation method with parameters of $\kappa = 3.2/r_{\text{cut}}$ and $K_{\text{max}} = \kappa L_{\text{box}} + 1$ is used for simulations with Coulomb interactions [219, 44]. For simulations in the NpT single-phase and NVT-Gibbs ensembles,[284, 92, 93] the pressure is calculated every 10 MC cycles (MCCs, with 1 MCC consisting of N randomly selected MC steps). The maximum displacement values for translational, rotational, and volume moves are adjusted during the equilibration period to give an average acceptance rate of 40%, and they are fixed during the efficiency tests. Simulations using the COM cutoff, linked-cell list, and k-d tree yield the exact same trajectories as those using the complete-loop method.

For each test system, at least four independent simulations are performed on an Intel Haswell E5-2680v4 processor (Linux operating system) and *Intel Fortran Compiler* (version 13.0.1; Intel Corporation; Santa Clara, CA) with "-free", "-fpp", "-O3", "-xHost", and "-prec-div" flags. Each simulation is run for at least one hour to minimize the effect of overhead time during startup and to provide a meaningful average because the

time for moves is highly variable (due to move termination before completion of range search). One CPU core is used for each run. The timing is recorded by calling the CPU timer at the start and the end of the simulation. The timing data are confirmed by using the profiler *Allinea MAP* (version 5.0; Allinea Software; San Jose, CA) for a few selected systems. Timing results are reported here as CPU time per 10^3 or 10^4 MC steps (MCSs) and reflect the average taken over the complete runs. Statistical uncertainties determined from the independent simulations are reported as 95% confidence interval.

5.3 Results and Discussion

5.3.1 Lennard-Jonesium in the NVT ensemble

The first system investigated is Lennard-Jonesium in the canonical (NVT) ensemble [98] close to its triple point with reduced temperature $T^* = k_\mathrm{B}T/\epsilon = 0.72$, reduced density $\rho^* = \rho\sigma^3 = 0.75$, and reduced cutoff $r^*_\mathrm{cut} = r_\mathrm{cut}/\sigma = 4$. Figure 5.4 illustrates CPU timings as a function of the number of molecules. In each translation attempt, the energies associated with the displaced particle at its current and its trial position are computed. Thus, the CPU time per MCS for the complete loop method scales as $\mathcal{O}(N)$. For the linked-cell list method, there is no performance improvement when $L_\mathrm{box}/r_\mathrm{cut} < 4$, since the number of cells examined in the range search is equal to the total number of cells. When $L_\mathrm{box}/r_\mathrm{cut} \geq 4$, the use of linked-cell list can dramatically boost the efficiency by reducing the number of distance calculations. As the box length crosses an integer multiple of r_cut, more cells are added to the system, resulting in an enhancement and discontinuity in the performance. When the system size is large enough, the timing converges to a constant value; this matches the asymptotic scaling expected for the linked-cell list method ($\mathcal{O}(1)$).

The performance of the k-d tree method fits the theoretical performance scaling of

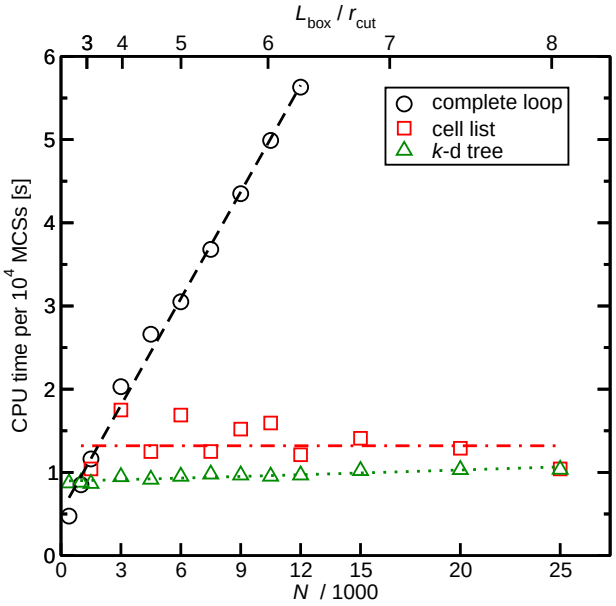

Figure 5.4: CPU timings for the Lennard-Jonesium simulations in the NVT ensemble ($T^* = 0.72$, $\rho^* = 0.75$, and $r^*_{\text{cut}} = 4$) as a function of the number of particles, N. Dashed, dash-dotted, and dotted lines are the theoretical performance scaling for the complete loop ($\mathcal{O}(N)$), linked-cell list ($\mathcal{O}(1)$), and k-d tree ($\mathcal{O}(\log_2 N)$) approaches, respectively. The alternate x-axis shows the box dimension relative to r_{cut}.

$\mathcal{O}(\log_2 N)$ fairly well. Compared to the linked-cell list, there is a 20-80% performance enhancement resulting from the use of k-d tree when $3000 \leq N \leq 20000$. This is likely due to a larger reduction of the number of distance calculations in the range search. A linked-cell list range search covers a total volume of $27r^3_{\text{cut}}$ which is approximately six times greater than the minimum volume that needs to be examined $\frac{4}{3}\pi r^3_{\text{cut}}$. In contrast, the k-d tree range search requires only a few extra distance calculations for particles outside of the spherical cutoff. When the system size is over a certain threshold (in this case $N > 1000$), the benefit achieved by reducing the number of distance calculations outweighs the extra computational cost of maintaining the tree structure, thus resulting in a performance enhancement for the k-d tree over the complete loop and linked-cell

list approaches. When $N \geq 25000$, the linked-cell list becomes the method of choice because of its better asymptotic scaling ($\mathcal{O}(1)$ versus $\mathcal{O}(\log_2 N)$). Nevertheless, the use of the k-d tree data structure can increase the performance in most practical cases ($1000 < N < 25000$) for Lennard-Jones particle simulations of a single phase.

5.3.2 Lennard-Jonesium in the NVT-Gibbs ensemble

Simulations of Lennard-Jones particles in the NVT-Gibbs ensemble close to the triple point are used to illustrate the performance of the k-d tree on a variety of other MC moves. In these Gibbs ensemble simulations the probability for selecting a volume move is set to $2/N$, that for a swap move is adjusted to yield one accepted move per MCC, and the remainder is taken up by translational moves. The timing results for Lennard-Jonesium in the Gibbs ensemble are shown in Figure 5.5. As explained previously, the linked-cell list implemented in MCCCS-MN can only be applied to one simulation box, and therefore a fair comparison between linked-cell list and k-d tree is made by using these approaches only for the liquid box and the complete loop method for the vapor box.

For the volume moves, use of the k-d tree method for both boxes with scaling as $\mathcal{O}(N\log_2 N)$ for individual boxes can increase the efficiency by a factor ranging from 1.2 to 3.3 for $N = 3000$ and 12000, respectively, compared to the complete loop method ($\mathcal{O}(N^2)$). It should be noted that distribution of the particles over two simulation boxes leads to the presence of lower order terms that somewhat obscure the scaling behavior. The timings for the linked-cell list applied to only the liquid box exhibit slightly oscillatory behavior due to discontinuous changes in the number of cells, and their values are close to those obtained from applying the k-d tree method to only the liquid phase in all of the system sizes studied here. Applying the k-d tree method to both boxes during a volume move versus only to the liquid phase yields a slight performance

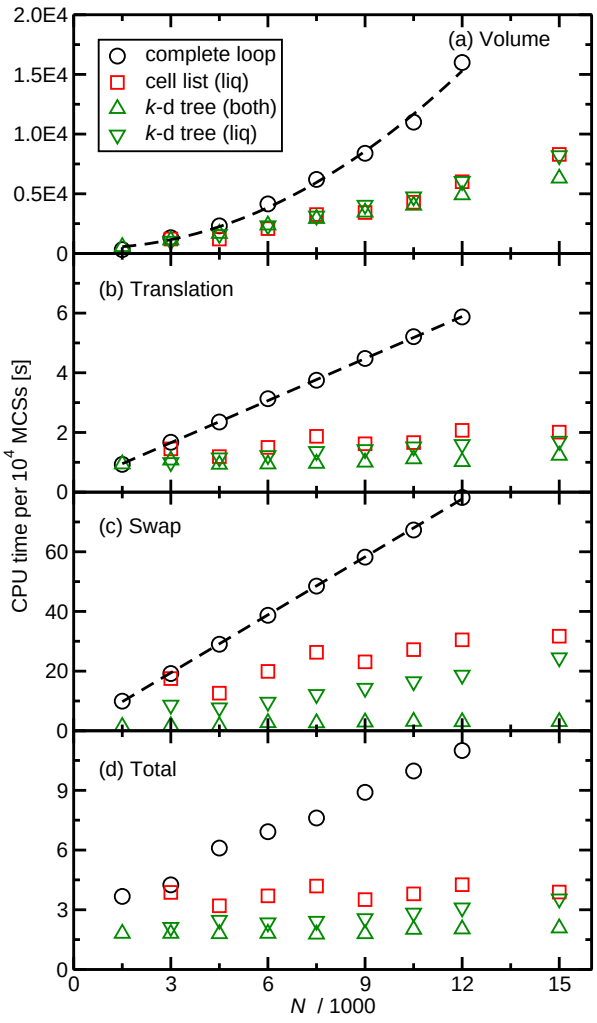

Figure 5.5: CPU timings for Lennard-Jonesium simulations in the NVT Gibbs ensemble ($T^* = 0.72$, $r^*_{\rm cut} = 4$, and 20% of the molecules in the vapor phase) for (a) volume moves, (b) translational moves, (c) swap moves, and (d) total set of moves as functions of the number of molecules, N. Dashed lines represent the scaling for the complete loop method (quadratic for the volume moves, and linear for translational and swap moves).

boost when $N > 6000$.

The performance data for translational moves in the Gibbs ensemble mirror those in the NVT ensemble when considering that about 80% and 20% of the particles are found in the liquid- and vapor-phase boxes, respectively, i.e., the performance gain is a weighted average. Compared to the complete loop method, the k-d tree applied to both boxes yields a speed-up by factors of 1.5 to 5.8 for $N = 3000$ and 12000, respectively. The average efficiency increases over the linked-cell list method observed by applying the k-d tree to both or only to the liquid phase are 1.6 and 1.2, respectively.

Applying the k-d tree ($\mathcal{O}(\log_2 N)$) to both phases can significantly boost efficiency for the particle swap moves by a factor of 12 for a 3000-particle system compared to the complete loop approach ($\mathcal{O}(N)$), and this efficiency gap widens as system size increases. Compared to the linked-cell list, the use of the k-d tree applied to only the liquid phase can provide an average efficiency improvement of 80% for the system size range investigated here, and this average gain increases to a factor of 9 when the k-d tree is applied to both phases. The dramatic efficiency improvement from the use of the k-d tree for swap moves can be attributed to two aspects: First, there are more energy calculations per coordinate update for a swap move compared to a translation move (requiring energy calculations for 32 current and trial sites in a swap move versus two sites in a translational move), because the k-d tree speeds up the energy calculations at the cost of slower coordinate updates; Second, a smaller r_{cut} value is used (5 Å in this case) is used for the trial sites energy calculations, and the k-d tree is more efficient when the ratio $L_{\text{box}}/r_{\text{cut}}$ is large.

The overall timing of such a Gibbs ensemble simulation for Lennard-Jonesium with a reasonable distribution of MC moves ($2/N$ volume moves, one accepted swap move per MCC, and the remainder for translational moves) shown in Figure 5.5(d) indicates that there is a two- to six-fold increase in efficiency compared to the complete loop method

(when the k-d tree is applied to both boxes). The efficiency increase compared to the linked-cell list method ranges from a factor of 1.8 for $N = 3000$ to a factor of 1.1 for $N = 15000$ (when both methods are applied only to the liquid box), and these gains increase to factors of 2.1 and 1.9 when the k-d tree is used for both boxes. Therefore, use of the k-d tree is recommended for Lennard-Jones particle simulations in the NVT-Gibbs ensemble for system sizes as small as 1500 to as large as 15000 particles.

5.3.3 n-Butane in the NpT ensemble

In addition to Lennard-Jonesium, the effect of using the k-d tree data structure on the simulation efficiency is also studied for a few different types of molecules that are represented by models with multiple interaction sites. The TraPPE–UA force field [95] is used here to simulate n-butane molecules in the NpT ensemble ($T = 255$ K $\approx 0.6\,T_{\mathrm{crit}}$, $p = 1$ bar). The COM cutoff method (i.e., complete loop over molecules instead of over sites) is used as control for the range search because it is the state-of-the-art approach for simulations of multiple-bead molecules [107]. For example, using a complete loop over sites for translational and rotational moves is slower by factors of 6 and 12 for $N = 1500$ and 6000, respectively, than applying the COM cutoff. The performance gain from the COM cutoff would be even larger for volume and CBMC moves (the former scaling as $\mathcal{O}(N^2)$ and the latter benefitting from the shorter inner cutoff). Since the reduced temperature is relatively low, 50% of the MCSs are dedicated to CBMC moves, the probability for volume moves is set to $2/N$, and the remainder is equally distributed between translational and rotational moves. The performance data for the n-butane simulations using a standard cutoff of 14 Å are depicted in Figure 5.6. For the volume moves, use of the k-d tree is found to be slower by factors of 2.4 and 1.3 for $N = 1500$ and 6000, respectively, than using the COM cutoff. This ratio increases to 2.4 compared to the linked-cell list for $N = 6000$. Among the three methods, the linked-cell list method

is the most efficient one when $N > 3000$ because of its computational complexity $\mathcal{O}(N)$. Although the CPU requirements for the volume moves using the k-d tree method scale as $\mathcal{O}(N\log_2 N)$ versus a scaling of $\mathcal{O}(N^2)$ for the COM cutoff, the latter benefits from the 16-fold reduction in distance calculations for molecules that satisfy eq 5.1.

For translational, rotational, and CBMC conformational moves, the COM cutoff, linked-cell list, and k-d tree methods scale as $\mathcal{O}(N)$, $\mathcal{O}(1)$, and $\mathcal{O}(\log_2 N)$, respectively. However, different ratios of energy calculations to coordinate updates are associated with these move types and the performance data differ markedly. For translational and rotational moves, the linked-cell list is found to be most efficient for the range of system sizes investigated here. It should be noted that accepted rotations around the COM do not require any update in the linked-cell list. Due to the better scaling behavior of the k-d tree method, it becomes more efficient than the COM cutoff for $N = 6000$, but it is on average a factor of 1.8 less efficient than the linked-cell list. The failure of the k-d tree method to provide any efficiency gain over the COM cutoff when applied to translational and rotational moves for typical system sizes will be discussed later together with results from simulations of water and ethanol. Nevertheless, for the CBMC conformational moves, the k-d tree is the most efficient method and yields an average boost of 12% compared to the linked-cell list and up to a factor of 1.7 ($N = 6000$) compared to the COM cutoff. Similar to the CBMC swap move, a CBMC conformational move also involves computing energies associated with multiple trial positions during the regrowth. Hence, the efficiency increase is mainly due to a larger number of energy calculations (with a smaller r_{cut}) per coordinate update in contrast to translational and rotational moves.

When considering the timings for the entire move set, then all three methods pose a similar CPU demand for $N = 1500$. For $N \geq 3000$, both k-d tree and linked-cell methods offer significant benefits over the COM cutoff, but the k-d tree is found to be

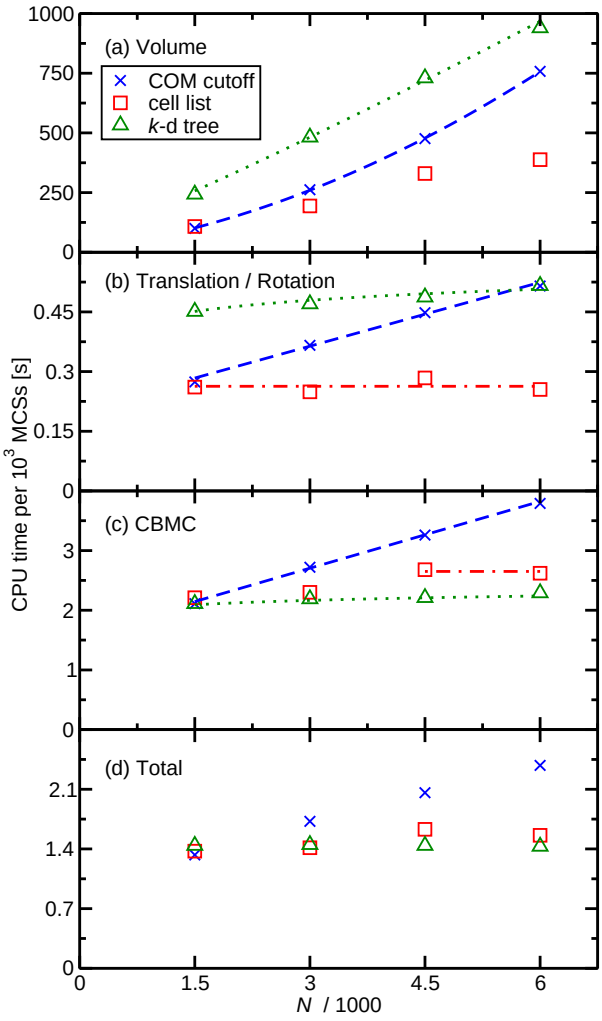

Figure 5.6: CPU timings for simulations of TraPPE–UA n-butane in the NpT ensemble ($T = 255$ K, $p = 1$ bar) for (a) volume moves, (b) translational and rotational moves, (c) CBMC conformational moves, and (d) total set of moves as functions of the number of molecules, N. Dashed, dashed-dotted, and dotted lines indicate the asymptotic scaling for the COM cutoff, linked-cell list, and k-d tree approaches, respectively.

slightly more efficient for $N = 4500$ and 6000. Therefore, the use of the k-d tree is recommended for flexible molecules without partial charges like n-butane for a sufficient large system size ($N = 3000$ for n-butane at this condition in the NpT ensemble).

5.3.4 Ethanol in the NpT ensemble

The applicability of the k-d tree data structure to a molecular system with Coulomb interactions is probed by NpT simulations for ethanol represented by the TraPPE–UA force field [216] ($T = 323$ K $\approx 0.63\,T_{\mathrm{crit}}$, $p = 1$ bar). The TraPPE–UA ethanol model consists of four interaction sites, i.e., the same number as TraPPE–UA butane. For ethanol, however, three of the four sites (all but the methyl group) carry partial charges and three of the four sites (all but the hydroxyl hydrogen atom) interact via Lennard-Jones potentials. Since CBMC conformational moves have a lower acceptance rate for ethanol than for n-butane, the fraction of CBMC moves is increased to 60%, the probability for volume moves is set to $2/N$, and the remainder is equally distributed between translational and rotational moves. The performance data for the ethanol simulations are shown in Figure 5.7.

The performances for all move types exhibit significant differences between ethanol and n-butane because of the contribution from the reciprocal space part of the Ewald summation in the energy calculation. The Ewald summation as implemented in MCCCS–MN scales as $\mathcal{O}(N^2)$ for volume moves, and $\mathcal{O}(N)$ for translational, rotational, and CBMC moves, and it increases the computational complexity for all range search algorithms. Overall, the absolute differences between the three methods are similar between ethanol and n-butane. For example, at $N = 6000$, the time differences for the total set of moves between the k-d tree and COM cutoff method are 0.94 and 0.95 seconds per 10^3 MCSs for ethanol and n-butane, respectively. Due to the additional expense of the reciprocal space part of the Ewald summation, the relative efficiency enhancements from

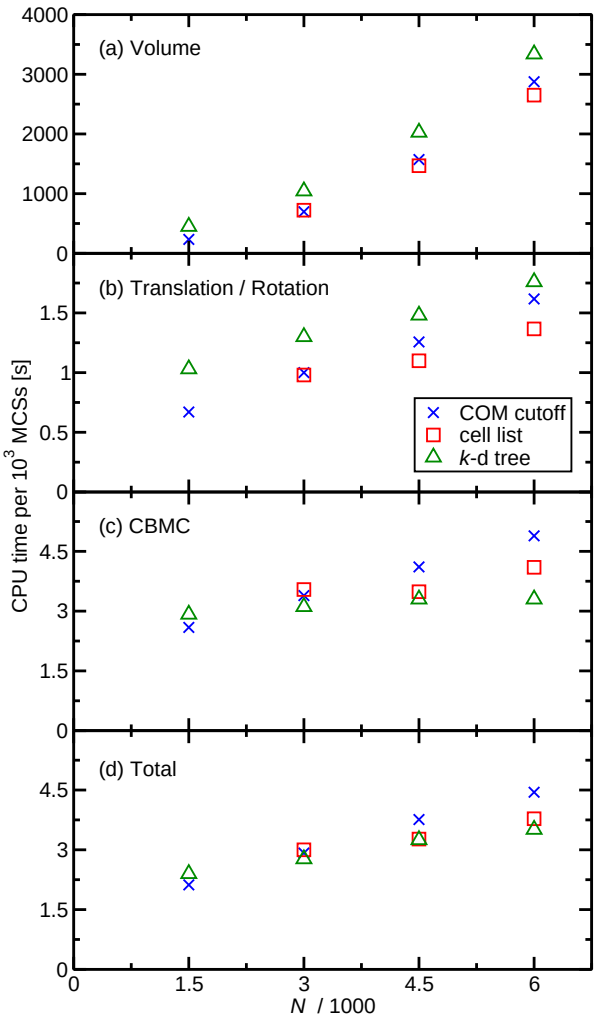

Figure 5.7: CPU timings for simulations of TraPPE–UA ethanol in the NpT ensemble ($T = 323$ K, $p = 1$ bar) for (a) volume moves, (b) translational and rotational moves, (c) CBMC conformational moves, and (d) total set of moves as functions of the number of molecules, N.

use of the k-d tree or linked-cell list compared to the COM cutoff are much smaller for ethanol. Nevertheless, the trends are similar to those for n-butane simulations: Applying the k-d tree data structure is not as efficient as the COM cutoff or linked-cell list methods for volume, translational, and rotational moves, but the k-d tree yields efficiency enhancements of 12 and 20% over the linked-cell list for CBMC move when $N = 3000$ and 6000, respectively. In conclusion, for ethanol, the COM cutoff performs best for the total move set when $N \leq 1500$, but use of the k-d tree brings overall efficiency gains of around 8% compared to the linked-cell list for $3000 \leq N \leq 6000$.

5.3.5 TIP4P water in the NpT and NVT-Gibbs ensemble

For simulations of n-butane and ethanol, the efficiency increase from using the k-d tree is mostly dominated by the contribution from the efficiency enhancement for CBMC conformational moves. TIP4P water [266], a rigid molecule, is studied to assess whether MC simulations can benefit from the use of the k-d tree without CBMC conformational moves. Results for simulations in the single-phase NpT ensemble ($T = 298$ K $\approx 0.46\,T_{\text{crit}}$, $p = 1$ bar) and two-phase NVT-Gibbs ensemble ($T = 298$ K) are shown in Figure 5.8 and Figure 5.9, respectively. It is noteworthy that for the NpT ensemble simulations, the use of the linked-cell list does not provide any performance gain over the COM cutoff because the ratio $L_{\text{box}}/L_{\text{cell}}$ is always below the threshold value of 4 for the system sizes investigated here. Note that if a smaller r_{cut} value is used, the linked-cell list method can provide some efficiency improvement. For the NVT-Gibbs ensemble simulations, the liquid box size is too small to apply the linked-cell list. Similar to the simulations of n-butane and ethanol, using the k-d tree method decreases the efficiency for volume, translational, and rotational moves compared to the COM cutoff approach by a factor close to 2 for $N = 2000$ in the NpT and Gibbs ensembles. The efficiency gap shrinks as the system size increases but, at $N = 8000$, using the k-d tree still results

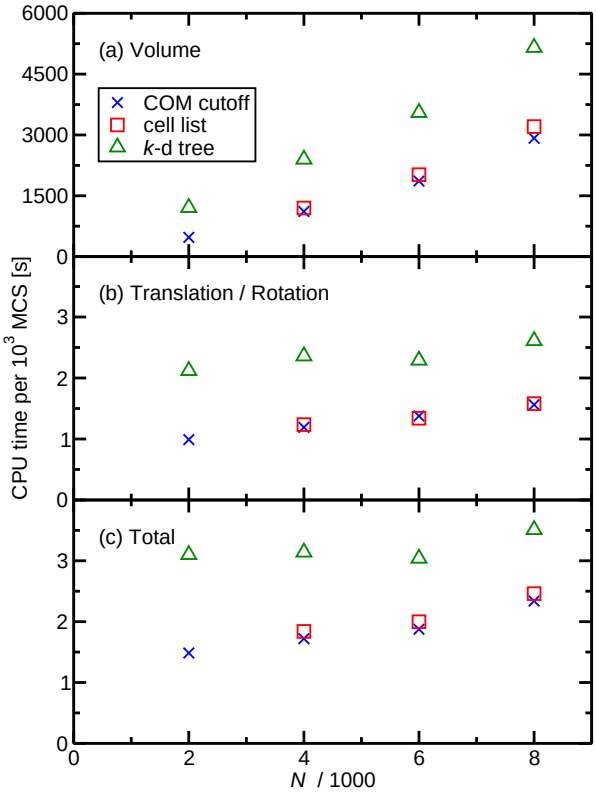

Figure 5.8: CPU timings for simulations of TIP4P water in the NpT ensemble ($T = 298$ K, $p = 1$ bar) for (a) volume moves, (b) translational or rotational moves, and (c) total set of moves as functions of the number of molecules, N.

in an efficiency decrease by a factor of 1.7 for these move types. For the total move set ($2/N$ dedicated to volume moves, and the remainder of moves equally distributed between translations and rotations), applying the k-d tree decreases efficiency by factors of 2.1 and 1.5 for $N = 2000$ and 8000, respectively. Thus, using the k-d tree cannot provide any speed-up for TIP4P water in the NpT ensemble for all of the system sizes covered in this study.

In the NVT-Gibbs ensemble, the swap moves are dramatically sped up by using the k-d tree, similar to the swap-move efficiency gain for Lennard-Jonesium in Gibbs

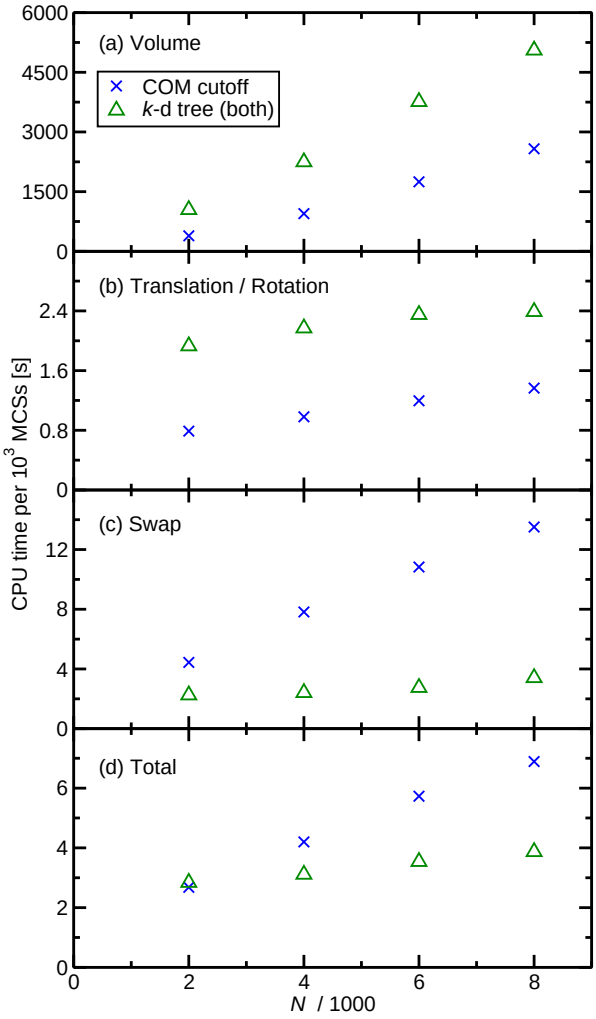

Figure 5.9: CPU timings for simulations of TIP4P water in the NVT-Gibbs ensemble (T = 298 K, 20% of the molecules in the vapor phase) for (a) volume moves, (b) translation or rotation moves, (c) swap moves, and (d) total set of moves as functions of the number of molecules, N.

ensemble. At $N = 2000$, the k-d tree can result in a factor of 2 increase in the swap-move efficiency. As the system size increases, the benefit of using the k-d tree is even more pronounced, reaching a boost of 4 for swap moves at $N = 8000$. With the significant efficiency gain for the swap move, applying the k-d tree for the total set of moves ($2/N$ volume moves, 40% swap moves, and the remainder equally distributed among translations and rotations) yields a speed-up by factors of 1.4 and 1.8 for $N = 4000$ and 8000, respectively, whereas it is slightly less efficient for $N = 2000$. Therefore, using the k-d tree is not recommended for simulations of rigid water models in the NpT ensemble, but it can significantly speed up simulations in the NVT-Gibbs ensemble when $N > 2000$.

5.3.6 Performance cross-comparison between different molecules

To further understand the failure of the k-d tree data structure to increase the efficiency for translational and rotational moves in systems composed of small molecules (n-butane, ethanol, and water), we define an efficiency enhancement (EE) factor as the ratio of the CPU time needed for simulations with only translational moves using either the complete loop or COM cutoff methods over that using the k-d tree method. As can be seen from Figure 5.10, the EE factors with the complete loop as control are larger than unity for all but three systems; the exceptions are values of 0.55, 0.96, and 0.95 for Lennard-Jonesium with $N = 400$ and 1000 and water with $N = 2000$ ($N_{\text{site}} = 8000$). Furthermore, the EE factors with the complete loop as control agree fairly well for n-butane and Lennard-Jonesium systems. The smaller EE factors observed for ethanol and water systems are due to the expense for calculating the reciprocal space part of the Ewald summation that does not benefit from the k-d tree data structure.

When the COM cutoff method is used as the control, then the EE factors for the three different molecules represented by four interaction sites are mostly smaller than

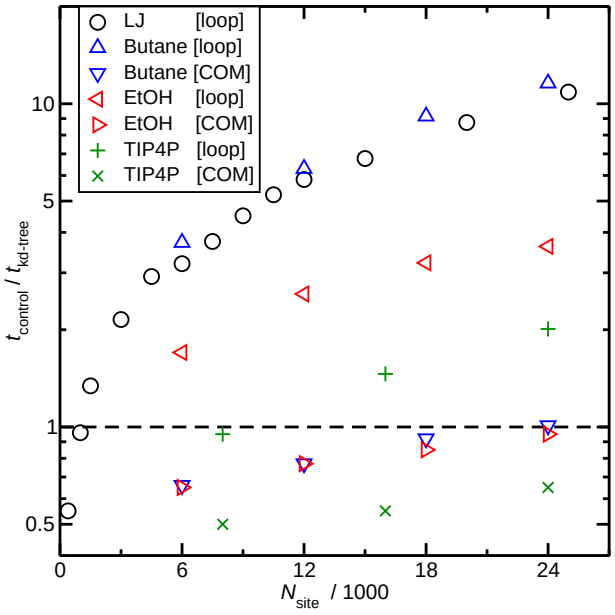

Figure 5.10: The efficiency enhancement factors as functions of the total number of interaction sites, N_{site}. The black circles represent data for Lennard-Jonesium. The cyan, magenta, and green symbols show data for TraPPE–UA *n*-butane, TraPPE–UA ethanol (EtOH), and TIP4P water, where up/left triangles and pluses denote enhancement factors calculated using the complete loop, and right/down triangles and crosses denote the enhancement factors computed using the COM cutoff as control method.

unity (the exception is a value of 1.01 for the system with 6000 *n*-butane molecules, i.e., 24000 interaction sites). Clearly, the COM cutoff is very efficient for these small molecules and speeds up the simulations for *n*-butane by factors ranging from 5.6 to 11 for systems with 1500 to 6000 molecules; with further increases in the system size, this value should converge to 16 because one COM distance calculation replaces 16 site–site distance calculations. In contrast, the gain from replacing the complete loop with the COM cutoff reaches only a value of about 3 for systems with 6000 ethanol and water molecules ($N_{site} = 24000$). Overall it is evident that the somewhat disappointing EE factors for *n*-butane, ethanol, and water compared to the COM cutoff method can

mainly be attributed to their small molecular sizes, where the COM cutoff method can provide the largest enhancements (see eq 5.1). In addition, the three partial charges in the ethanol and water models further reduce the relative benefit of using the k-d tree data structure.

5.3.7 36-Carbon propylene oligomer in the NpT ensemble

It can be deduced from the preceding discussion that using the k-d tree may result in improved efficiency, even for translational and rotational moves, for long-chain molecules because the COM cutoff provides a smaller efficiency gain when r_{intra} is large (see eq 5.1). This hypothesis is verified by simulations for 2,4,6,8,10,12,14,16,18,20,22-undecamethylpentacosane (a 36-carbon oligomer of propylene [121]) in the NpT ensemble ($T = 440$ K, $p = 1$ bar). Due to the large number of intramolecular degrees of freedom, the probability for CBMC conformational moves is set to 80%; the fraction of volume moves is $2/N$, and the remainder is equally distributed between translations and rotations. Note that the COM-based linked-cell list cannot be used for the current system sizes since the large value of Δr_{intra} would require a minimum cell size larger than twice the site–site cutoff (i.e., $L_{\text{box}}/r_{\text{cut}} \approx 5.5$ for $N = 400$).

CPU timings for systems consisting of 100 to 400 molecules are shown in Figure 5.11. The better scaling for the k-d tree method results in a widening gap with increasing system size for volume moves, and the gain in efficiency reaches a factor of 1.5 for $N = 400$. For translational, rotational, and CBMC conformational moves, the CPU time saturates as the system size reaches $N = 200$ for both COM cutoff and k-d tree approaches. Nevertheless, using the k-d tree boosts efficiency for translational and rotational moves by factors of 1.4 and 2.2 for $N = 100$ and $N \geq 200$, respectively, whereas the speed-up obtained for CBMC moves is only a factor of 1.5. The smaller efficiency increase for CBMC moves is due to a smaller inner cutoff during the calculation of the Rosenbluth

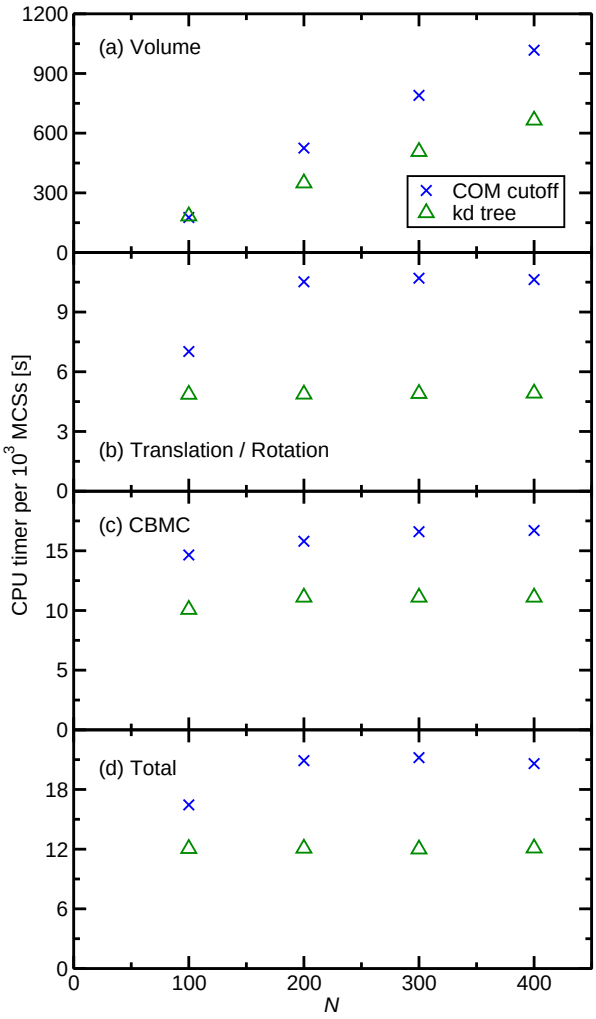

Figure 5.11: CPU timings for simulations of 2,4,6,8,10,12,14,16,18,20,22-undecamethylpentacosane in the NpT ensemble (T = 440 K, p = 1 bar) for (a) volume moves, (b) translational or rotational moves, (c) CBMC conformational moves, and (d) total set of moves as functions of the number of molecules, N.

weights [107]. For the total set of moves, the gain in efficiency due to use of the k-d tree ranges from 1.4 to 1.7 for $100 \leq N \leq 400$. This example proves that polymer simulations (especially those for models without partial charges, e.g. polyolefins) can benefit significantly from applying the k-d tree data structure, even for a relative small system size. For comparison, the simulation for 4500 n-butane molecules ($N_{\text{site}} = 18000$) yields an EE factor of 1.4 with the COM method as control, whereas the same factor is already obtained for 100 $C_{36}H_{74}$ molecules ($N_{\text{site}} = 3600$). To summarize, the k-d tree is recommended for the 36-carbon propylene oligomer simulation in the NpT ensemble for all the system sizes investigated here.

5.4 Conclusions

In this work, the benefits from applying the k-d tree data structure with its $\mathcal{O}(\log_2 N)$ scaling for the range search are assessed for MC simulations. To this extent, the k-d tree data structure is implemented in the in-house MCCCS–MN simulation software package. Simulations are performed for a variety of molecular systems and ensembles to provide a comprehensive efficiency comparison.

The efficiency tests lead to a few rules of thumb on cases where using the k-d tree can provide significant efficiency increases. First, (semi-)flexible molecules that require CBMC conformational moves benefit more than rigid molecules with the same number of interaction sites. Second, Gibbs ensemble (or grand canonical ensemble) simulations where CBMC swap moves are used, benefit more than simulations in closed ensembles. Third, single-site or long-chain molecules where use of the COM cutoff provides little efficiency gain, benefit more than those for compact multi-site molecules. Fourth, for molecules that include partial charges, the percentage increase in efficiency is smaller but the absolute decrease in CPU time is the same compared to molecules that have the same number of interaction sites but do not contain partial charges.

The largest increases in efficiency are observed here for Lennard-Jones particle in the Gibbs ensemble and $C_{36}H_{74}$ (described by a model with 36 united atoms) that both satisfy three out of the four rules (i.e., the second, third, and fourth rules for the former and the first, third, and fourth rules for the latter simulations). For these two cases, the k-d tree is the most efficient method even for the smallest system size considered here ($N = 1500$ for Lennard-Jonesium and $N = 100$ for $C_{36}H_{74}$).

The simulations for n-butane in the NpT ensemble satisfy only the first and fourth rules, those for ethanol in the NpT ensemble satisfy only the first rule, and those for water in the Gibbs ensemble satisfy only the second rule. Thus, an efficiency increase from application of the k-d tree data structure is only observed for system sizes exceeding 3000, 4500, and 4000 molecules, respectively. The simulations for water in the NpT ensemble, however, do not satisfy any of the four rules, and the COM cutoff approach is the fastest approach for $2000 \leq N \leq 8000$ (for even larger systems, the linked-cell list would become most efficient).

The four rules encompass a wide range of molecules and ensembles and applying the k-d data structure enables significant efficiency gains for simulations covering practical system sizes from about 3000 to 30000 interaction sites. Thus, implementation of the k-d tree data structure is recommended for MC simulations of molecular systems that satisfy at least one of the four rules.

Chapter 6

Conclusions and Future Work

In this thesis, particle based simulations were used in conjunction with accurate molecular models to compute phase behavior of polymer blends and block polymers, which enables predictive materials design for polymeric materials.

For polymer blends, Gibbs ensemble Monte Carlo simulations were utilized to calculate the mixing thermodynamics for the oligomeric blends. Then, theoretical framework developed in the thesis was used to extrapolate it to polymer blends with arbitrary molecular weight of interest. The effect of molecular weight distribution and dispersity of the polymers on the phase diagram was also probed by simulations. Results from this thesis allow for the accurate determination of phase diagram for binary polymer blends.

For block polymers, I leveraged molecular dynamics simulations to probe their self-assembled microstructures. It was demonstrated that one can screen candidate molecules for desired structures and properties using simulations. Furthermore, the molecular level insight gained from simulations, such as chain conformation and hydrogen bonding capabilities, gives rise to better understanding of the phase behavior, and thus, better molecular design.

Future work extending this thesis research includes (i) developing molecular models

to extend the current method to other polymers, (ii) investigating phase behavior of ternary polymer mixtures, and (iii) refining the underlying thermodynamic theory with the aid of simulation data.

Bibliography

[1] P. S. Chum and K. W. Swogger, Olefin Polymer Technologies: History and Recent Progress at the Dow Chemical Company, *Prog. Polym. Sci.*, **2008**, *33*, 797–819.

[2] H. S. Jeon, J. H. Lee, and N. P. Balsara, Predictions of the Thermodynamic Properties of Multicomponent Polyolefin Blends from Measurements on Two-Component Systems, *Macromolecules*, **1998**, *31*, 3328–3339.

[3] C. M. Bates and F. S. Bates, 50th Anniversary Perspective: Block Polymers-Pure Potential, *Macromolecules*, **2017**, *50*, 3–22.

[4] K. F. Freed and J. Dudowicz; Influence of Monomer Molecular Structure on the Miscibility of Polymer Blends; In *Phase Behaviour of Polymer Blends*, pages 63–126. Springer, 2005.

[5] L. Bai, M. Li, and C. Nan, Preparation and Characterization of Polymer Electrolyte Based on Comb-Shaped Poly (oxyethylene) with Ether Side Chains, *Key Eng. Mater.*, **2012**, *519*, 14–19.

[6] S. N. Patel, A. E. Javier, G. M. Stone, S. A. Mullin, and N. P. Balsara, Simultaneous Conduction of Electronic Charge and Lithium Ions in Block Copolymers, *ACS nano*, **2012**, *6*, 1589–1600.

[7] S. A. Chopade, J. G. Au, Z. Li, P. W. Schmidt, M. A. Hillmyer, and T. P. Lodge, Robust Polymer Electrolyte Membranes with High Ambient-Temperature Lithium-Ion Conductivity via Polymerization-Induced Microphase Separation, *ACS Appl. Mater. Interfaces*, **2017**, *9*, 14561–14565.

[8] S. Huang, L. Bai, M. Trifkovic, X. Cheng, and C. W. Macosko, Controlling the Morphology of Immiscible Cocontinuous Polymer Blends via Silica Nanoparticles Jammed at the Interface, *Macromolecules*, **2016**, *49*, 3911–3918.

[9] L. Bai, R. Sharma, X. Cheng, and C. W. Macosko, Kinetic Control of Graphene Localization in Co-Continuous Polymer Blends via Melt Compounding, *Langmuir*, **2018**, *34*, 1073–1083.

[10] L. Bai, Controlling Cocontinuous Polymer Blends with Nanofillers Jammed at the Interface. Ph.D. Thesis, University of Minnesota, 2017.

[11] J. Xu, V. Mittal, and F. S. Bates, Toughened Isotactic Polypropylene: Phase Behavior and Mechanical Properties of Blends with Strategically Designed Random Copolymer Modifiers, *Macromolecules*, **2016**, *49*, 6497–6506.

[12] J. M. Eagan, J. Xu, R. Di Girolamo, C. M. Thurber, C. W. Macosko, A. M. LaPointe, F. S. Bates, and G. W. Coates, Combining Polyethylene and Polypropylene: Enhanced Performance with PE/iPP Multiblock Polymers, *Science*, **2017**, *355*, 814–816.

[13] T. Li, J. Zhang, D. K. Schneiderman, L. F. Francis, and F. S. Bates, Toughening Glassy Poly(lactide) with Block Copolymer Micelles, *ACS Macro Lett.*, **2016**, *5*, 359–364.

[14] J. Xu, M. J. Howard, V. Mittal, and F. S. Bates, Block Copolymer Micelle Toughened Isotactic Polypropylene, *Macromolecules*, **2017**, *50*, 6421–6432.

[15] M. T. Savoji, D. Zhao, R. J. Muisener, K. Schimossek, K. Schoeller, T. P. Lodge, and M. A. Hillmyer, Poly (alkyl methacrylate)-Grafted Polyolefins as Viscosity Modifiers for Engine Oil: A New Mechanism for Improved Performance, *Ind. Eng. Chem. Res.*, **2018**.

[16] Y. Zhang, Q. Chen, J. Ge, and Z. Liu, Controlled Display of Enzyme Activity with a Stretchable Hydrogel, *Chem. Commun.*, **2013**, *49*, 9815–9817.

[17] Y. Jiang, D. Sprouse, J. E. Laaser, Y. Dhande, T. M. Reineke, and T. P. Lodge, Complexation of Linear DNA and Poly (styrenesulfonate) with Cationic Copolymer Micelles: Effect of Polyanion Flexibility, *J. Phys. Chem. B*, **2017**, *121*, 6708–6720.

[18] Z. Li, L. M. Johnson, R. G. Ricarte, L. J. Yao, M. A. Hillmyer, F. S. Bates, and T. P. Lodge, Enhanced Performance of Blended Polymer Excipients in Delivering a Hydrophobic Drug through the Synergistic Action of Micelles and HPMCAS, *Langmuir*, **2017**, *33*, 2837–2848.

[19] S. Zhi, J. Xu, R. Deng, L. Wan, and Z. Xu, Poly (vinylidene fluoride) Ultrafiltration Membranes Containing Hybrid Silica Nanoparticles: Preparation, Characterization and Performance, *Polymer*, **2014**, *55*, 1333–1340.

[20] J. R. Werber, C. O. Osuji, and M. Elimelech, Materials for Next-Generation Desalination and Water Purification Membranes, *Nat. Rev. Mater.*, **2016**, *1*, 16018.

[21] M. Luo and T. H. Epps III, Directed Block Copolymer Thin Film Self-Assembly: Emerging Trends in Nanopattern Fabrication, *Macromolecules*, **2013**, *46*, 7567–7579.

[22] C. Sinturel, F. S. Bates, and M. A. Hillmyer, High χ-Low N Block Polymer: How Far can We Go?, *ACS Macro Lett.*, **2015**, *4*, 1044–1050.

[23] F. S. Bates, M. A. Hillmyer, T. P. Lodge, C. M. Bates, K. T. Delaney, and G. H. Fredrickson, Multiblock Polymers: Panacea or Pandora's Box?, *Science*, **2012**, *336*, 434–440.

[24] P. J. Flory, Thermodynamics of High Polymer Solutions, *J. Chem. Phys.*, **1942**, *10*, 51–61.

[25] M. L. Huggins, Solutions of Long Chain Compounds, *J. Chem. Phys.*, **1941**, *9*, 440–440.

[26] R. Krishnamoorti, W. W. Graessley, G. T. Dee, D. J. Walsh, L. J. Fetters, and D. J. Lohse, Pure Component Properties and Mixing Behavior in Polyolefin Blends, *Macromolecules*, **1996**, *29*, 367–376.

[27] R. Krishnamoorti, W. W. Graessley, L. J. Fetters, R. T. Garner, and D. J. Lohse, Anomalous Mixing Behavior of Polyisobutylene with Other Polyolefins, *Macromolecules*, **1995**, *28*, 1252–1259.

[28] C. C. Han, B. J. Bauer, J. C. Clark, Y. Muroga, Y. Matsushita, M. Okada, Q. Trancong, T. Chang, and I. C. Sanchez, Temperature, Composition and Molecular-Weight Dependence of the Binary Interaction Parameter of Polystyrene/poly (vinyl methyl ether) Blends, *Polymer*, **1988**, *29*, 2002–2014.

[29] H. Petri and B. A. Wolf, Composition-Dependent Flory-Huggins Parameters: Molecular Weight Influences at High Concentrations, *Macromol. Chem. Phys.*, **1995**, *196*, 2321–2333.

[30] R. Krishnamoorti, W. W. Graessley, N. P. Balsara, and D. J. Lohse, The Compositional Dependence of Thermodynamic Interactions in Blends of Model Polyolefins, *J. Chem. Phys.*, **1994**, *100*, 3894–3904.

[31] A. J. Nedoma, M. L. Robertson, N. S. Wanakule, and N. P. Balsara, Measurements of the Composition and Molecular Weight Dependence of the Flory–Huggins Interaction Parameter, *Macromolecules*, **2008**, *41*, 5773–5779.

[32] S. K. Kumar, Chemical Potentials of Polymer Blends from Monte Carlo Simulations: Consequences on SANS-Determined. Chi Parameters, *Macromolecules*, **1994**, *27*, 260–271.

[33] T. H. Russell, B. J. Edwards, and B. Khomami, Characterization of the Flory-Huggins Interaction Parameter of Polymer Thermodynamics, *Europhys. Lett.*, **2015**, *108*, 66003.

[34] J. K. Maranas, M. Mondello, G. S. Grest, S. K. Kumar, P. G. Debenedetti, and W. W. Graessley, Liquid Structure, Thermodynamics, and Mixing Behavior of Saturated Hydrocarbon Polymers. 1. Cohesive Energy Density and Internal Pressure, *Macromolecules*, **1998**, *31*, 6991–6997.

[35] J. K. Maranas, S. K. Kumar, P. G. Debenedetti, W. W. Graessley, M. Mondello, and G. S. Grest, Liquid Structure, Thermodynamics, and Mixing Behavior of Saturated Hydrocarbon Polymers. 2. Pair Distribution Functions and the Regularity of Mixing, *Macromolecules*, **1998**, *31*, 6998–7002.

[36] W. W. Graessley, R. Krishnamoorti, G. C. Reichart, N. P. Balsara, L. J. Fetters, and D. J. Lohse, Regular and Irregular Mixing in Blends of Saturated Hydrocarbon Polymers, *Macromolecules*, **1995**, *28*, 1260–1270.

[37] F. S. Bates, M. F. Schulz, J. H. Rosedale, and K. Almdal, Correlation of Binary Polyolefin Phase Behavior with Statistical Segment Length Asymmetry, *Macromolecules*, **1992**, *25*, 5547–5550.

[38] G. H. Fredrickson, A. J. Liu, and F. S. Bates, Entropic Corrections to the Flory-Huggins Theory of Polymer Blends: Architectural and Conformational Effects, *Macromolecules*, **1994**, *27*, 2503–2511.

[39] K. S. Schweizer, Analytic PRISM Theory of Structurally Asymmetric Polymer Blends and Copolymers, *Macromolecules*, **1993**, *26*, 6050–6067.

[40] K. S. Schweizer and C. Singh, Microscopic Solubility-Parameter Theory of Polymer Blends: General Predictions, *Macromolecules*, **1995**, *28*, 2063–2080.

[41] K. F. Freed and J. Dudowicz, Influence of Short Chain Branching on the Miscibility of Binary Polymer Blends: Application to Polyolefin Mixtures, *Macromolecules*, **1996**, *29*, 625–636.

[42] K. F. Freed, J. Dudowicz, and K. W. Foreman, Molecular Mechanisms for Disparate Miscibilities of Poly(propylene) and Head-to-Head Poly(propylene) with Other Polyolefins, *J. Chem. Phys.*, **1998**, *108*, 7881–7886.

[43] R. P. White and J. E. G. Lipson, Free Volume, Cohesive Energy Density, and Internal Pressure as Predictors of Polymer Miscibility, *Macromolecules*, **2014**, *47*, 3959–3968.

[44] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications;* Academic Press, 2001.

[45] B. Widom, Some Topics in the Theory of Fluids, *J. Chem. Phys.*, **1963**, *39*, 2808–2812.

[46] S. T. Cui, H. D. Cochran, P. T. Cummings, and S. K. Kumar, Computer Simulations of the Static Scattering from Model Polymer Blends, *Macromolecules*, **1997**, *30*, 3375–3382.

[47] M. Pütz, J. G. Curro, and G. S. Grest, Self-Consistent Integral Equation Theory for Polyolefins: Comparison to Molecular Dynamics Simulations and X-Ray Scattering, *J. Chem. Phys.*, **2001**, *114*, 2847–2860.

[48] P. De Gennes, *Scaling Concepts in Polymer Physics;* Cornell University Press, 1979.

[49] D. Heine, D. T. Wu, J. G. Curro, and G. S. Grest, Role of Intramolecular Energy on Polyolefin Miscibility: Isotactic Polypropylene/polyethylene Blends, *J. Chem. Phys.*, **2003**, *118*, 914–924.

[50] E. Jaramillo, D. T. Wu, G. S. Grest, and J. G. Curro, Anomalous Mixing Behavior of Polyisobutylene/polypropylene Blends: Molecular Dynamics Simulation Study, *J. Chem. Phys.*, **2004**, *120*, 8883–8886.

[51] J. McCarty and M. G. Guenza, Multiscale Modeling of Binary Polymer Mixtures: Scale Bridging in the Athermal and Thermal Regime, *J. Chem. Phys.*, **2010**, *133*, 094904.

[52] G. Miquelard-Garnier and S. Roland, Beware of the Flory Parameter to Characterize Polymer-Polymer Interactions: A Critical Reexamination of the Experimental Literature, *Eur. Polym. J.*, **2016**, *84*, 111–124.

[53] J. I. Siepmann, A Method for the Direct Calculation of Chemical Potentials for Dense Chain Systems, *Mol. Phys.*, **1990**, *70*, 1145–1158.

[54] S. K. Kumar, I. Szleifer, and A. Z. Panagiotopoulos, Determination of the Chemical Potentials of Polymeric Systems from Monte Carlo Simulations, *Phys. Rev. Lett.*, **1991**, *66*, 2935.

[55] G. Mooij, D. Frenkel, and B. Smit, Direct Simulation of Phase-Equilibria of Chain Molecules, *J. Phys.: Condens. Matter*, **1992**, *4*, L255.

[56] J. J. de Pablo, M. Laso, and U. W. Suter, Estimation of the Chemical Potential of Chain Molecules by Simulation, *J. Chem. Phys.*, **1992**, *96*, 6157–6162.

[57] B. Chen and J. I. Siepmann, Partitioning of Alkane and Alcohol Solutes between Water and (Dry or Wet) 1-octanol, *J. Am. Chem. Soc.*, **2000**, *122*, 6464–6467.

[58] T. W. Rosch and J. R. Errington, Investigation of the Phase Behavior of an Embedded Charge Protein Model through Molecular Simulation, *J. Phys. Chem. B*, **2007**, *111*, 12591–12598.

[59] W. Shi and E. J. Maginn, Continuous Fractional Component Monte Carlo: An Adaptive Biasing Method for Open System Atomistic Simulations, *J. Chem. Theory Comput.*, **2007**, *3*, 1451–1463.

[60] B. Yoo, E. Marin-Rimoldi, R. G. Mullen, A. Jusufi, and E. J. Maginn, Discrete Fractional Component Monte Carlo Simulation Study of Dilute Nonionic Surfactants at the Air–Water Interface, *Langmuir*, **2017**, *33*, 9793–9802.

[61] P. Bai and J. I. Siepmann, Gibbs Ensemble Monte Carlo Simulations for the Liquid–Liquid Phase Equilibria of Dipropylene Glycol Dimethyl Ether and Water: A Preliminary Report, *Fluid Phase Equilib.*, **2011**, *310*, 11–18.

[62] D. B. Harwood, C. J. Peters, and J. I. Siepmann, A Monte Carlo Simulation Study of the Liquid–Liquid Equilibria for Binary Dodecane/ethanol and Ternary Dodecane/ethanol/water Mixtures, *Fluid Phase Equilib.*, **2016**, *407*, 269–279.

[63] W. Zhang, E. D. Gomez, and S. T. Milner, Predicting Flory-Huggins χ from Simulations, *Phys. Rev. Lett.*, **2017**, *119*, 017801.

[64] N. C. Karayiannis, V. G. Mavrantzas, and D. N. Theodorou, A Novel Monte Carlo Scheme for the Rapid Equilibration of Atomistic Model Polymer Systems of Precisely Defined Molecular Architecture, *Phys. Rev. Lett.*, **2002**, *88*, 105503.

[65] C. D. Wick, J. I. Siepmann, and D. N. Theodorou, Microscopic Origins for the Favorable Solvation of Carbonate Ether Copolymers in CO_2, *J. Am. Chem. Soc.*, **2005**, *127*, 12338–12342.

[66] Q. Wang, R. F. Nealey, and J. J. de Pablo, Monte Carlo Simulations of Asymmetric Diblock Copolymer Thin Films Confined between Two Homogeneous Surfaces, *Macromolecules*, **2001**, *34*, 3458–3470.

[67] A. Arora, D. C. Morse, F. S. Bates, and K. D. Dorfman, Commensurability and Finite Size Effects in Lattice Simulations of Diblock Copolymers, *Soft Matter*, **2015**, *11*, 4862–4867.

[68] T. M. Gillard, P. Medapuram, D. C. Morse, and F. S. Bates, Fluctuations, Phase Transitions, and Latent Heat in Short Diblock Copolymers: Comparison of Experiment, Simulation, and Theory, *Macromolecules*, **2015**, *48*, 2801–2811.

[69] S. Mao, Q. MacPherson, and A. J. Spakowitz, Fluctuation Effects in Semiflexible Diblock Copolymers, *ACS Macro Lett.*, **2017**, *7*, 59–64.

[70] G. J. Papakonstantopoulos, K. C. Daoulas, M. Muller, and J. J. de Pablo, Monte Carlo Simulation Study of Diblock Copolymer Self Assembly, *arXiv preprint arXiv:1604.05265*, **2016**.

[71] A. Chremos, A. Nikoubashman, and A. Z. Panagiotopoulos, Flory-Huggins Parameter χ, from Binary Mixtures of Lennard-Jones Particles to Block Copolymer Melts, *J. Chem. Phys.*, **2014**, *140*, 054909.

[72] J. H. Hildebrand and R. L. Scott, *The Solubility of Nonelectrolytes;* Dover: New York, 1964.

[73] T. Lindvig, M. L. Michelsen, and G. M. Kontogeorgis, A Flory–Huggins Model Based on the Hansen Solubility Parameters, *Fluid Phase Equilib.*, **2002**, *203*, 247–260.

[74] N. Rai, A. J. Wagner, R. B. Ross, and J. I. Siepmann, Application of the TraPPE Force Field for Predicting the Hildebrand Solubility Parameters of Organic Solvents and Monomer Units, *J. Chem. Theory Comput.*, **2008**, *4*, 136–144.

[75] N. P. Balsara, L. J. Fetters, N. Hadjichristidis, D. J. Lohse, C. C. Han, W. W. Graessley, and R. Krishnamoorti, Thermodynamic Interactions in Model Polyolefin Blends Obtained by Small-Angle Neutron Scattering, *Macromolecules*, **1992**, *25*, 6137–6147.

[76] W. W. Graessley, R. Krishnamoorti, N. P. Balsara, R. J. Butera, L. J. Fetters, D. J. Lohse, D. N. Schulz, and J. A. Sissano, Thermodynamics of Mixing for Blends of Model Ethylene-Butene Copolymers, *Macromolecules*, **1994**, *27*, 3896–3901.

[77] R. Krishnamoorti, W. W. Graessley, N. P. Balsara, and D. J. Lohse, Structural Origin of Thermodynamic Interactions in Blends of Saturated Hydrocarbon Polymers, *Macromolecules*, **1994**, *27*, 3073–3081.

[78] G. C. Reichart, W. W. Graessley, R. A. Register, R. Krishnamoorti, and D. J. Lohse, Anomalous Attractive Interactions in Polypropylene Blends, *Macromolecules*, **1997**, *30*, 3036–3041.

[79] A. Neelakantan, R. Stine, and J. K. Maranas, Chain Packing in Ethylene-Butene Copolymers, *Macromolecules*, **2003**, *36*, 3721–3731.

[80] A. Neelakantan and J. K. Maranas, Intra- and Intermolecular Packing in Polyolefin Blends, *Macromolecules*, **2004**, *37*, 8473–8480.

[81] J. D. Weinhold, S. K. Kumar, C. Singh, and K. S. Schweizer, Athermal Stiffness Blends: A Comparison of Monte Carlo Simulations and Integral Equation Theory, *J. Chem. Phys.*, **1995**, *103*, 9460–9474.

[82] C. Singh and K. S. Schweizer, Correlation Effects and Entropy-Driven Phase Separation in Athermal Polymer Blends, *J. Chem. Phys.*, **1995**, *103*, 5814–5832.

[83] C. Singh and K. S. Schweizer, Molecular Theory of the Miscibility of Hydrocarbon Blends, *Macromolecules*, **1995**, *28*, 8692–8695.

[84] C. Singh and K. S. Schweizer, Coupled Enthalpic-Packing Effects on the Miscibility of Conformationally Asymmetric Polymer Blends, *Macromolecules*, **1997**, *30*, 1490–1508.

[85] K. W. Foreman, K. F. Freed, and I. M. Ngola, Nonrandom Mixing in Polymer Blends: Implications for Phase Behavior, *J. Chem. Phys.*, **1997**, *107*, 4688–4704.

[86] J. E. G. Lipson and R. P. White, Connecting Theory and Experiment to Understand Miscibility in Polymer and Small Molecule Mixture, *J. Chem. Eng. Data*, **2014**, *59*, 3289–3300.

[87] R. P. White, J. E. G. Lipson, and J. S. Higgins, How Pure Components Control Polymer Blend Miscibility, *Macromolecules*, **2012**, *45*, 8861–8871.

[88] R. P. White, J. E. G. Lipson, and J. S. Higgins, New Correlations in Polymer Blend Miscibility, *Macromolecules*, **2012**, *45*, 1076–1084.

[89] N. Rai, J. I. Siepmann, N. E. Schultz, and R. B. Ross, Pressure Dependence of the Hildebrand Solubility Parameter and the Internal Pressure: Monte Carlo

Simulations for External Pressures Up to 300 MPa, *J. Phys. Chem. C*, **2007**, *111*, 15634–15641.

[90] J. McCarty, A. J. Clark, J. Copperman, and M. G. Guenza, An Analytical Coarse-Graining Method which Preserves the Free Energy, Structural Correlations, and Thermodynamic State of Polymer Melts from the Atomistic to the Mesoscale, *J. Chem. Phys.*, **2014**, *140*, 204913.

[91] T. C. Clancy, M. Pütz, J. D. Weinhold, J. G. Curro, and W. L. Mattice, Mixing of Isotactic and Syndiotactic Polypropylenes in the Melt, *Macromolecules*, **2000**, *33*, 9452–9463.

[92] A. Z. Panagiotopoulos, Direct Determination of Phase Coexistence Properties of Fluids by Monte Carlo Simulation in a New Ensemble, *Mol. Phys.*, **1987**, *61*, 813–826.

[93] A. Z. Panagiotopoulos, N. Quirke, M. Stapleton, and D. J. Tildesley, Phase Equilibria by Simulation in the Gibbs Ensemble: Alternative Derivation, Generalization and Application to Mixture and Membrane Equilibria, *Mol. Phys.*, **1988**, *63*, 527–545.

[94] J. I. Siepmann, S. Karaborni, and B. Smit, Simulating the Critical Properties of Complex Fluids, *Nature*, **1993**, *365*, 330–332.

[95] M. G. Martin and J. I. Siepmann, Transferable Potentials for Phase Equilibria. 1. United-Atom Description of *n*-Alkanes, *J. Phys. Chem. B*, **1998**, *102*, 2569–2577.

[96] M. G. Martin and J. I. Siepmann, Novel Configurational-Bias Monte Carlo Method for Branched Molecules. Transferable Potentials for Phase Equilibria. 2. United-Atom Description of Branched Alkanes, *J. Phys. Chem. B*, **1999**, *103*, 4508–4517.

[97] G. C. Maitland, M. Rigby, E. Smith, W. Wakeham, and D. Henderson, *Intermolecular Forces: Their Origin and Determination;* Oxford University Press, 1981.

[98] W. W. Wood and F. R. Parker, Monte Carlo Equation of State of Molecules Interacting with the Lennard-Jones Potential. I. A Supercritical Isotherm at about Twice the Critical Temperature, *J. Chem. Phys.*, **1957**, *27*, 720–733.

[99] J. I. Siepmann *et al.*, Monte Carlo for Complex Chemical Systems–Minnesota. Version 17.1; Minneapolis, MN, 2017.

[100] J. J. De Pablo and J. M. Prausnitz, Phase Equilibria for Fluid Mixtures from Monte-Carlo Simulation, *Fluid Phase Equilib.*, **1989**, *53*, 177–189.

[101] J. I. Siepmann and I. R. McDonald, Monte Carlo Simulations of Mixed Monolayers, *Mol. Phys.*, **1992**, *75*, 255–259.

[102] M. G. Martin and J. I. Siepmann, Predicting Multicomponent Phase Equilibria and Free Energies of Transfer for Alkanes by Molecular Simulation, *J. Am. Chem. Soc.*, **1997**, *119*, 8921–8924.

[103] J. W. Teh, A. Rudin, and J. C. Keung, A Review of Polyethylene–Polypropylene Blends and Their Compatibilization, *Adv. Polym. Technol.*, **1994**, *13*, 1–23.

[104] R. Thomann, J. Kressler, S. Setz, C. Wang, and R. Mülhaupt, Morphology and Phase Behaviour of Blends of Syndiotactic and Isotactic Polypropylene. 1. X-ray Scattering, Light Microscopy, Atomic Force Microscopy, and Scanning Electron Microscopy, *Polymer*, **1996**, *37*, 2627–2634.

[105] P. Bai and J. I. Siepmann, Selective Adsorption from Dilute Solutions: Gibbs Ensemble Monte Carlo Simulations, *Fluid Phase Equilib.*, **2013**, *351*, 1–6.

[106] C. D. Wick and J. I. Siepmann, Self-Adapting Fixed-End-Point Configurational-Bias Monte Carlo Method for the Regrowth of Interior Segments of Chain Molecules with Strong Intramolecular Interactions, *Macromolecules*, **2000**, *33*, 7207–7218.

[107] T. J. H. Vlugt, M. G. Martin, B. Smit, J. I. Siepmann, and R. Krishna, Improving the Efficiency of the Configurational-Bias Monte Carlo Algorithm, *Mol. Phys.*, **1998**, *94*, 727–733.

[108] M. Lagache, P. Ungerer, A. Boutin, and A. H. Fuchs, Prediction of Thermodynamic Derivative Properties of Fluids by Monte Carlo Simulation, *Phys. Chem. Chem. Phys.*, **2001**, *3*, 4333–4339.

[109] C. M. Colina, C. G. Olivera-Fuentes, F. R. Siperstein, M. Lisal, and K. E. Gubbins, Thermal Properties of Supercritical Carbon Dioxide by Monte Carlo Simulations, *Mol. Simul.*, **2003**, *29*, 405–412.

[110] B. D. Smith and R. Srivastava, *Thermodynamic Data for Pure Compounds. Part A: Hydrocarbons and Ketones;* Elsevier Science Limited: New York, 1986.

[111] E. Von Meerwall, S. Beckman, J. Jang, and W. L. Mattice, Diffusion of Liquid *n*-Alkanes: Free-Volume and Density Effects, *J. Chem. Phys.*, **1998**, *108*, 4299–4304.

[112] T. E. Daubert, Vapor-Liquid Critical Properties of Elements and Compounds. 5. Branched Alkanes and Cycloalkanes, *J. Chem. Eng. Data*, **1996**, *41*, 365–372.

[113] B. Chen and J. I. Siepmann, Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, *J. Phys. Chem. B*, **1999**, *103*, 5370–5379.

[114] C. Berro, F. Laichoubi, and E. Rauzy, Isothermal Vapor-Liquid Equilibria and Excess Volumes for the Systems *n*-Hexane + Ethylbenzene, 2-Methylpentane + *n*-Heptane, and 2-Methylpentane + *n*-Octane, *J. Chem. Eng. Data*, **1991**, *36*, 474–478.

[115] L. Bai, S. He, J. W. Fruehwirth, A. Stein, C. W. Macosko, and X. Cheng, Localizing Graphene at the Interface of Cocontinuous Polymer Blends: Morphology, Rheology, and Conductivity of Cocontinuous Conductive Polymer Composites, *J. Rheol.*, **2017**, *61*, 575–587.

[116] T. Li, M. J. Heinzer, L. F. Francis, and F. S. Bates, Engineering Superior Toughness in Commercially Viable Block Copolymer Modified Epoxy Resin, *J. Polym. Sci., Part B: Polym. Phys.*, **2016**, *54*, 189–204.

[117] K. Jin and J. M. Torkelson, Tg-Confinement Effects in Strongly Miscible Blends of Poly (2, 6-dimethyl-1, 4-phenylene oxide) and Polystyrene: Roles of Bulk Fragility and Chain Segregation, *Polymer*, **2017**, *118*, 85–96.

[118] M. T. Irwin, R. J. Hickey, S. Xie, F. S. Bates, and T. P. Lodge, Lithium Salt-Induced Microstructure and Ordering in Diblock Copolymer/Homopolymer Blends, *Macromolecules*, **2016**, *49*, 4839–4849.

[119] S. Xie and T. P. Lodge, Phase Behavior of Binary Polymer Blends Doped with Salt, *Macromolecules*, **2018**, *51*, 266–274.

[120] R. D. Groot and P. B. Warren, Dissipative Particle Dynamics: Bridging the Gap between Atomistic and Mesoscopic Simulation, *J. Chem. Phys.*, **1997**, *107*, 4423–4435.

[121] Q. P. Chen, J. D. Chu, R. F. DeJaco, T. P. Lodge, and J. I. Siepmann, Molecular Simulation of Olefin Oligomer Blend Phase Behavior, *Macromolecules*, **2016**, *49*, 3975–3985.

[122] A. J. Nedoma, M. L. Robertson, N. S. Wanakule, and N. P. Balsara, Measurements of the Flory–Huggins Interaction Parameter Using a Series of Critical Binary Blends, *Ind. Eng. Chem. Res.*, **2008**, *47*, 3551–3553.

[123] A. A. Teran and N. P. Balsara, Thermodynamics of Block Copolymers with and without Salt, *J. Phys. Chem. B*, **2013**, *118*, 4–17.

[124] D. C. Morse and J. K. Chung, On the Chain Length Dependence of Local Correlations in Polymer Melts and a Perturbation Theory of Symmetric Polymer Blends, *J. Chem. Phys.*, **2009**, *130*, 224901.

[125] P. Knychała, K. Timachova, M. Banaszak, and N. P. Balsara, 50th Anniversary Perspective: Phase Behavior of Polymer Solutions and Blends, *Macromolecules*, **2017**, *50*, 3051–3065.

[126] R. Kita, K. Kubota, and T. Dobashi, Static and Dynamic Light Scattering of a Critical Polydisperse Polymer Solution, *Phys. Rev. E*, **1998**, *58*, 793.

[127] R. Kita, T. Dobashi, T. Yamamoto, M. Nakata, and K. Kamide, Coexistence Curve of a Polydisperse Polymer Solution near the Critical Point, *Phys. Rev. E*, **1997**, *55*, 3159.

[128] M. Nakata and T. Dobashi, Coexistence Curve of Polystyrene in Methylcyclohexane. V. Critical Behavior of the Three-Phase Coexistence Curve, *J. Chem. Phys.*, **1986**, *84*, 5782–5786.

[129] M. Nakata, T. Dobashi, Y. Inakuma, and K. Yamamura, Coexistence Curve of Polystyrene in Methylcyclohexane. X. Two-Phase Coexistence Curves for Ternary Solutions Near the Tricritical Compositions, *J. Chem. Phys.*, **1999**, *111*, 6617–6624.

[130] W. Shen, G. R. Smith, C. M. Knobler, and R. L. Scott, Tricritical Phenomena in Bimodal Polymer Solutions: Three-Phase Coexistence Curves for the System Polystyrene (1)+ Polystyrene (2)+ Methylcyclohexane, *J. Phys. Chem.*, **1990**, *94*, 7943–7949.

[131] R. Koningsveld and A. J. Staverman, Liquid–Liquid Phase Separation in Multicomponent Polymer Solutions. I. Statement of the Problem and Description of Methods of Calculation, *J. Polym. Sci., Part A-2*, **1968**, *6*, 305–323.

[132] R. Koningsveld and A. J. Staverman, Liquid–Liquid Phase Separation in Multicomponent Polymer Solutions. III. Cloud-Point Curves, *J. Polym. Sci., Part B: Polym. Phys.*, **1968**, *6*, 349–366.

[133] K. Šolc, Cloud-Point Curves of Polymer Solutions, *Macromolecules*, **1970**, *3*, 665–673.

[134] J. A. Gualtieri, J. M. Kincaid, and G. Morrison, Phase Equilibria in Polydisperse Fluids, *J. Chem. Phys.*, **1982**, *77*, 521–536.

[135] C. Qian, S. J. Mumby, and B. E. Eichinger, Phase Diagrams of Binary Polymer Solutions and Blends, *Macromolecules*, **1991**, *24*, 1655–1661.

[136] Y. Hu, X. Ying, D. T. Wu, and J. M. Prausnitz, Liquid-Liquid Equilibria for Solutions of Polydisperse Polymers. Continuous Thermodynamics for the Close-Packed Lattice Model, *Macromolecules*, **1993**, *26*, 6817–6823.

[137] Y. Hu, X. Ying, D. T. Wu, and J. M. Prausnitz, Continuous Thermodynamics for Polydisperse Polymer Solutions, *Fluid Phase Equilib.*, **1995**, *104*, 229–252.

[138] G. Ten Brinke and I. Szleifer, Liquid-Liquid Phase Separation in Polydisperse Polymer Solutions: The Distribution Coefficient, *Macromolecules*, **1995**, *28*, 5434–5439.

[139] C. Huang and M. Olvera de La Cruz, The Early Stages of the Phase Separation Dynamics in Polydisperse Polymer Blends, *Macromolecules*, **1994**, *27*, 4231–4241.

[140] R. S. Shresth, R. C. McDonald, and S. C. Greer, Molecular Weight Distributions of Polydisperse Polymers in Coexisting Liquid Phases, *J. Chem. Phys.*, **2002**, *117*, 9037–9049.

[141] R. M. L. Evans, D. J. Fairhurst, and W. C. K. Poon, Universal Law of Fractionation for Slightly Polydisperse Systems, *Phys. Rev. Lett.*, **1998**, *81*, 1326.

[142] H. Xu and M. Baus, Density Functional Theory of Phase Coexistence in Weakly Polydisperse Fluids, *Phys. Rev. E*, **2000**, *61*, 3249.

[143] M. R. Stapleton, D. J. Tildesley, and N. Quirke, Phase Equilibria in Polydisperse Fluids, *J. Chem. Phys.*, **1990**, *92*, 4456–4467.

[144] T. Kristóf and J. Liszi, Phase Coexistence and Critical Point Determination in Polydisperse Fluids, *Mol. Phys.*, **2001**, *99*, 167–173.

[145] N. B. Wilding, P. Sollich, M. Fasolo, and M. Buzzacchi, Phase Behavior and Particle Size Cutoff Effects in Polydisperse Fluids, *J. Chem. Phys.*, **2006**, *125*, 014908.

[146] L. Zhang and J. I. Siepmann, Pressure Dependence of the Vapor–Liquid–Liquid Phase Behavior in Ternary Mixtures Consisting of n-Alkanes, n-Perfluoroalkanes, and Carbon Dioxide, *J. Phys. Chem. B*, **2005**, *109*, 2911–2919.

[147] B. H. Zimm, The Scattering of Light and the Radial Distribution Function of High Polymer Solutions, *J. Chem. Phys.*, **1948**, *16*, 1093–1099.

[148] Q. P. Chen, B. Xue, and J. I. Siepmann, Using the k-d Tree Data Structure to Accelerate Monte Carlo Simulations, *J. Chem. Theory Comput.*, **2017**, *13*, 1556–1565.

[149] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids;* Clarendon Press Oxford, 1987.

[150] J. M. Stubbs, J. J. Potoff, and J. I. Siepmann, Transferable Potentials for Phase Equilibria. 6. United-Atom Description for Ethers, Glycols, Ketones, and Aldehydes, *J. Phys. Chem. B*, **2004**, *108*, 17596–17605.

[151] K. A. Kobe, A. E. Ravicz, and S. P. Vohra, Critical Properties and Vapor Pressures of Some Ethers and Heterocyclic Compounds, *J. Chem. Eng. Data*, **1956**, *1*, 50–56.

[152] W. V. Steele, R. D. Chirico, S. E. Knipmeyer, A. Nguyen, and N. K. Smith, Thermodynamic Properties and Ideal-Gas Enthalpies of Formation for Butyl Vinyl Ether, 1, 2-Dimethoxyethane, Methyl Glycolate, Bicyclo [2.2. 1] Hept-2-Ene, 5-Vinylbicyclo [2.2. 1] Hept-2-Ene, Trans-Azobenzene, Butyl Acrylate, Di-Tert-Butyl Ether, and Hexane-1, 6-Diol, *J. Chem. Eng. Data*, **1996**, *41*, 1285–1302.

[153] E. Brunner, Solubility of Hydrogen in Diols and Their Ethers, *J. Chem. Thermo.*, **1980**, *12*, 993–1002.

[154] S. K. Quadri, K. C. Khilar, A. P. Kudchadker, and M. J. Patni, Measurement of the Critical Temperatures and Critical Pressures of Some Thermally Stable or Mildly Unstable Esters, *J. Chem. Thermo.*, **1991**, *23*(1), 67–76.

[155] L. Sun, J. I. Siepmann, and M. R. Schure, Conformation and Solvation Structure for An Isolated *n*-Octadecane Chain in Water, Methanol, and Their Mixtures, *J. Phys. Chem. B*, **2006**, *110*, 10519–10525.

[156] T. Treszczanowicz and D. Cieślak, (Liquid + liquid) Equilibria in (a Dimethyl Ether of a Polyethene Glycol + an *n*-Alkane), *J. Chem. Thermo.*, **1993**, *25*, 661–665.

[157] N. R. Washburn, T. P. Lodge, and F. S. Bates, Ternary Polymer Blends as Model Surfactant Systems, *J. Phys. Chem. B*, **2000**, *104*, 6987–6997.

[158] X. Du, Z. Sun, and L. An, Effect of Chain-Length Dependence of Interaction Parameter on Spinodals for Polydisperse Polymer Blends, *Macromol. Theory Simul.*, **2006**, *15*, 440–445.

[159] J. L. Chen, B. Xue, D. B. Harwood, Q. P. Chen, C. J. Peters, and J. I. Siepmann, A Monte Carlo Simulation Study of the Interfacial Tension for Water/oil Mixtures at Elevated Temperatures and Pressures: Water/*n*-dodecane, Water/toluene, and Water/(*n*-dodecane + toluene), *Fluid Phase Equilib.*, **2017**.

[160] K. Šolc, Multiphase Equilibria in Solutions of Polydisperse Homopolymers. I. Three-Phase Equilibria in Ternary and Quasiternary Systems, *J. Polym. Sci. Poly,. Phys. Ed.*, **1982**, *20*, 1947–1961.

[161] C. C. Hsu and J. M. Prausnitz, Thermodynamics of Polymer Compatibility in Ternary Systems, *Macromolecules*, **1974**, *7*, 320–324.

[162] A. Ben-Naim and Y. Marcus, Solvation Thermodynamics of Nonionic Solutes, *J. Chem. Phys.*, **1984**, *81*, 2016–2027.

[163] M. Kurata, *Thermodynamics of Polymer Solutions;* CRC Press, 1982.

[164] P. Medapuram, J. Glaser, and D. C. Morse, Universal Phenomenology of Symmetric Diblock Copolymers near the Order–Disorder Transition, *Macromolecules*, **2015**, *48*, 819–839.

[165] M. R. Radlauer, C. Sinturel, Y. Asai, A. Arora, F. S. Bates, K. D. Dorfman, and M. A. Hillmyer, Morphological Consequences of Frustration in ABC Triblock Polymers, *Macromolecules*, **2016**, *50*, 446–458.

[166] D. P. Sanders, Advances in Patterning Materials for 193 nm Immersion Lithography, *Chem. Rev.*, **2010**, *110*, 321–360.

[167] C. M. Bates, M. J. Maher, D. W. Janes, C. J. Ellison, and C. G. Willson, Block Copolymer Lithography, *Macromolecules*, **2013**, *47*, 2–12.

[168] F. S. Bates and G. H. Fredrickson, Block Copolymers–Designer Soft Materials, *Phys. Today*, **1999**, *52*, 32–38.

[169] M. A. Hillmyer, F. S. Bates, K. Almdal, K. Mortensen, A. J. Ryan, and J. P. A. Fairclough, Complex Phase Behavior in Solvent-Free Nonionic Surfactants, *Science*, **1996**, *271*, 976.

[170] S. H. Kim, M. J. Misner, T. Xu, M. Kimura, and T. P. Russell, Highly Oriented and Ordered Arrays from Block Copolymers via Solvent Evaporation, *Adv. Mater.*, **2004**, *16*, 226–231.

[171] Y. S. Jung and C. A. Ross, Orientation-Controlled Self-Assembled Nanolithography Using a Polystyrene–Polydimethylsiloxane Block Copolymer, *Nano Lett.*, **2007**, *7*, 2046–2050.

[172] J. Bang, S. H. Kim, E. Drockenmuller, M. J. Misner, T. P. Russell, and C. J. Hawker, Defect-Free Nanoporous Thin Films from ABC Triblock Copolymers, *J. Am. Chem. Soc.*, **2006**, *128*, 7622–7629.

[173] C. Sinturel, M. Vayer, M. Morris, and M. A. Hillmyer, Solvent Vapor Annealing of Block Polymer Thin Films, *Macromolecules*, **2013**, *46*, 5399–5415.

[174] W. Li, P. F. Nealey, J. J. de Pablo, and M. Müller, Defect Removal in the Course of Directed Self-Assembly is Facilitated in the Vicinity of the Order-Disorder Transition, *Phys. Rev. Lett.*, **2014**, *113*, 168301.

[175] G. Blachut, Stephen M. Sirard, M. J. Maher, Y. Asano, Y. Someya, A. P. Lane, W. J. Durand, C. M. Bates, A. M. Dinhobl, R. Gronheid, and Willson C. G., A Hybrid Chemo-/Grapho-Epitaxial Alignment Strategy for Defect Reduction in Sub-10 nm Directed Self-Assembly of Silicon-Containing Block Copolymers, *Chem. Mater.*, **2016**, *28*, 8951–8961.

[176] C. Simão, W. Khunsin, N. Kehagias, M. Salaun, M. Zelsmann, M. A. Morris, and C. M. S. Torres, Order Quantification of Hexagonal Periodic Arrays Fabricated by In Situ Solvent-Assisted Nanoimprint Lithography of Block Copolymers, *Nanotechnology*, **2014**, *25*, 175703.

[177] W. J. Durand, M. C. Carlson, M. J. Maher, G. Blachut, L. J. Santos, S. Tein, V. Ganesan, C. J. Ellison, and C. G. Willson, Experimental and Modeling Study of Domain Orientation in Confined Block Copolymer Thin Films, *Macromolecules*, **2015**, *49*, 308–316.

[178] Q. Wang, Q. Yan, P. F. Nealey, and J. J. de Pablo, Monte Carlo Simulations of Diblock Copolymer Thin Films Confined between Two Homogeneous Surfaces, *J. Chem. Phys.*, **2000**, *112*, 450–464.

[179] C. M. Bates, T. Seshimo, M. J. Maher, W. J. Durand, J. D. Cushen, L. M. Dean, G. Blachut, C. J. Ellison, and C. G. Willson, Polarity-Switching Top Coats Enable Orientation of Sub–10-nm Block Copolymer Domains, *Science*, **2012**, *338*, 775–779.

[180] D. J. C. Herr, Directed Block Copolymer Self-Assembly for Nanoelectronics Fabrication, *J. Mater. Res.*, **2011**, *26*, 122–139.

[181] C. M. Bates and F. S. Bates, 50th Anniversary Perspective: Block Polymers–Pure Potential, *Macromolecules*, **2016**, *50*, 3–22.

[182] T. Xu, H. C. Kim, J. DeRouchey, C. Seney, C. Levesque, P. Martin, C. M. Stafford, and T. P. Russell, The Influence of Molecular Weight on Nanoporous Polymer Films, *Polymer*, **2001**, *42*, 9091–9095.

[183] Y. S. Jung, J. B. Chang, E. Verploegen, K. K. Berggren, and C. A. Ross, A Path to Ultranarrow Patterns Using Self-Assembled Lithography, *Nano Lett.*, **2010**, *10*, 1000–1005.

[184] J. W. Jeong, W. I. Park, M. Kim, C. A. Ross, and Y. S. Jung, Highly Tunable Self-Assembled Nanostructures from a Poly(2-vinylpyridine-*b*-dimethylsiloxane) Block Copolymer, *Nano Lett.*, **2011**, *11*, 4095–4101.

[185] L. M. Pitet, S. F. Wuister, E. Peeters, E. J. Kramer, C. J. Hawker, and E. W. Meijer, Well-Organized Dense Arrays of Nanodomains in Thin Films of Poly(dimethylsiloxane)-*b*-Poly(lactide) Diblock Copolymers, *Macromolecules*, **2013**, *46*, 8289–8295.

[186] J. D. Cushen, I. Otsuka, C. M. Bates, S. Halila, S. Fort, C. Rochas, J. A. Easley, E. L. Rausch, A. Thio, R. Borsali, C. G. Willson, and C. J. Ellison, Oligosaccharide/silicon-containing Block Copolymers with 5 nm Features for Lithographic Applications, *ACS Nano*, **2012**, *6*, 3424–3433.

[187] K. Aissou, I. Otsuka, C. Rochas, S. Fort, S. Halila, and R. Borsali, Nano-Organization of Amylose-*b*-Polystyrene Block Copolymer Films Doped with Bipyridine, *Langmuir*, **2011**, *27*, 4098–4103.

[188] W. J. Durand, G. Blachut, M. J. Maher, S. Sirard, S. Tein, M. C. Carlson, Y. Asano, S. X. Zhou, A. P. Lane, C. M. Bates, C. J. Ellison, and C. G. Willson, Design of High-χ Block Copolymers for Lithography, *J. Polym. Sci., Part A: Polym. Chem.*, **2015**, *53*, 344–352.

[189] S. Park, X. Liang, B. D. Harteneck, T. E. Pick, N. Hiroshiba, Y. Wu, B. A. Helms, and D. L. Olynick, Sub-10 nm Nanofabrication via Nanoimprint Directed Self-Assembly of Block Copolymers, *ACS Nano*, **2011**, *5*, 8523–8531.

[190] M. K. Mayeda, J. Hayat, T. H. Epps, and J. Lauterbach, Metal Oxide Arrays from Block Copolymer Thin Film Templates, *J. Mater. Chem. A*, **2015**, *3*, 7822–7829.

[191] R. H. Zha, B. F. M. de Waal, M. Lutz, A. J. P. Teunissen, and E. W. Meijer, End Groups of Functionalized Siloxane Oligomers Direct Block-Copolymeric or Liquid-Crystalline Self-Assembly Behavior, *J. Am. Chem. Soc.*, **2016**, *138*, 5693–5698.

[192] B. Van Genabeek, B. F. M. de Waal, M. M. J. Gosens, L. M. Pitet, A. R. A. Palmans, and E. W. Meijer, Synthesis and Self-Assembly of Discrete Dimethylsiloxane–Lactic Acid Diblock Co-Oligomers: The Dononacontamer and Its Shorter Homologues, *J. Am. Chem. Soc.*, **2016**, *138*, 4210–4218.

[193] W. Zha, C. D. Han, D. H. Lee, S. H. Han, J. K. Kim, J. H. Kang, and C. Park, Origin of the Difference in Order–Disorder Transition Temperature between Polystyrene-block-Poly(2-vinylpyridine) and Polystyrene-block-Poly(4-vinylpyridine) Copolymers, *Macromolecules*, **2007**, *40*, 2109–2119.

[194] J. G. Kennemur, M. A. Hillmyer, and F. S. Bates, Synthesis, Thermodynamics, and Dynamics of Poly(4-tert-butylstyrene-*b*-methyl methacrylate), *Macromolecules*, **2012**, *45*, 7228–7236.

[195] D. P. Sweat, M. Kim, S. R. Larson, J. W. Choi, Y. Choo, C. O. Osuji, and P. Gopalan, Rational Design of a Block Copolymer with a High Interaction Parameter, *Macromolecules*, **2014**, *47*, 6687–6696.

[196] J. G. Kennemur, L. Yao, F. S. Bates, and M. A. Hillmyer, Sub-5 nm Domains in Ordered Poly(cyclohexylethylene)-block-Poly(methyl methacrylate) Block Polymers for Lithography, *Macromolecules*, **2014**, *47*, 1411–1418.

[197] E. W. Cochran, D. C. Morse, and F. S. Bates, Design of ABC Triblock Copolymers near the ODT with the Random Phase Approximation, *Macromolecules*, **2003**, *36*, 782–792.

[198] A. S. Zalusky, R. Olayo-Valles, J. H. Wolf, and M. A. Hillmyer, Ordered Nanoporous Polymers from Polystyrene–Polylactide Block Copolymers, *J. Am. Chem. Soc.*, **2002**, *124*, 12761–12773.

[199] A. Baruth, M. Seo, C. H. Lin, K. Walster, A. Shankar, M. A. Hillmyer, and C. Leighton, Optimization of Long-Range Order in Solvent Vapor Annealed Poly(styrene)-block-Poly(lactide) Thin Films for Nanolithography, *ACS Appl. Mater. Interfaces*, **2014**, *6*, 13770–13781.

[200] I. Otsuka, Y. Zhang, T. Isono, C. Rochas, T. Kakuchi, T. Satoh, and R. Borsali, Sub-10 nm Scale Nanostructures in Self-Organized Linear Di-and Triblock Copolymers and Miktoarm Star Copolymers Consisting of Maltoheptaose and Polystyrene, *Macromolecules*, **2015**, *48*, 1509–1517.

[201] M. E. Vanderlaan and M. A. Hillmyer, "Uncontrolled" Preparation of Disperse Poly(lactide)-block-Poly(styrene)-block-Poly(lactide) for Nanopatterning Applications, *Macromolecules*, **2016**, *49*, 8031–8040.

[202] M. C. D. Carter, J. Jennings, F. W. Speetjens, D. M. Lynn, and M. K. Mahanthappa, A Reactive Platform Approach for the Rapid Synthesis and Discovery of High χ/Low N Block Polymers, *Macromolecules*, **2016**, *49*, 6268–6276.

[203] S. R. Nowak, W. Hwang, and L. R. Sita, Dynamic Sub-10-nm Nanostructured Ultrathin Films of Sugar–Polyolefin Conjugates Thermoresponsive at Physiological Temperatures, *J. Am. Chem. Soc.*, **2017**, *139*, 5281–5284.

[204] G. Jeong, D. M. Yu, J. K. D. Mapas, Z. Sun, J. Rzayev, and T. P. Russell, Realizing 5.4 nm Full Pitch Lamellar Microdomains by a Solid-State Transformation, *Macromolecules*, **2017**, *50*, 7148–7154.

[205] X. Feng, M. E. Tousley, M. G. Cowan, B. R. Wiesenauer, S. Nejati, Y. Choo, R. D. Noble, M. Elimelech, D. L. Gin, and C. O. Osuji, Scalable Fabrication of Polymer Membranes with Vertically Aligned 1 nm Pores by Magnetic Field Directed Self-Assembly, *ACS Nano*, **2014**, *8*, 11977–11986.

[206] X. Liu, S. Singh, E. L. Gibbemeyer, B. E. Tam, R. A. Urban, and B. R. Bakshi, The Carbon-Nitrogen Nexus of Transportation Fuels, *J. Cleaner Prod.*, **2018**, *180*, 790–803.

[207] X. Liu and B. R. Bakshi, Ecosystem Services in Life Cycle Assessment while Encouraging Techno-Ecological Synergies, *J. Ind. Ecol., In Press, DOI: 10.1111/jiec.12755.*

[208] S. Chandrasekhar, *Liquid Crystals;* Cambridge University, Cambridge, 1992.

[209] G. J. T. Tiddy, Surfactant-Water Liquid Crystal Phases, *Phys. Rep.*, **1980**, *57*, 1–46.

[210] M. D. Ward and M. J. Horner, Structure and Order in Soft Matter, *CrystEngComm*, **2004**, *6*, 401–407.

[211] M. Kikuchi and K. Binder, Microphase Separation in Thin Films of the Symmetric Diblock-Copolymer Melt, *J. Chem. Phys.*, **1994**, *101*, 3367–3377.

[212] R. G. Larson, Self-Assembly of Surfactant Liquid Crystalline Phases by Monte Carlo Simulation, *J. Chem. Phys.*, **1989**, *91*, 2479–2488.

[213] A. J. Peters, R. A. Lawson, B. D. Nation, P. J. Ludovice, and C. L. Henderson, Coarse-Grained Molecular Dynamics Modeling of the Kinetics of Lamellar Block Copolymer Defect Annealing, *J. Micro/Nanolithogr., MEMS, MOEMS*, **2016**, *15*, 013508–013508.

[214] G. Srinivas, D. E. Discher, and M. L. Klein, Self-Assembly and Properties of Diblock Copolymers by Coarse-Grain Molecular Dynamics, *Nat. Mater.*, **2004**, *3*, 638–644.

[215] W. B. Lee, R. Mezzenga, and G. H. Fredrickson, Anomalous Phase Sequences in Lyotropic Liquid Crystals, *Phys. Rev. Lett.*, **2007**, *99*, 187801.

[216] B. Chen, J. J. Potoff, and J. I. Siepmann, Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5.

United-Atom Description of Primary, Secondary, and Tertiary Alcohols, *J. Phys. Chem. B*, **2001**, *105*, 3093–3104.

[217] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc.*, **1996**, *118*, 11225–11236.

[218] F. Y. Jou, F. D. Otto, and A. E. Mather, The Solubility of Propane in 1, 2-Ethanediol at Elevated Pressures, *J. Chem. Thermo.*, **1993**, *25*, 37–40.

[219] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids;* Oxford University Press, 1989.

[220] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers, *SoftwareX*, **2015**, *1*, 19–25.

[221] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, GROMACS: Fast, Flexible, and Free, *J. Comput. Chem.*, **2005**, *26*, 1701–1718.

[222] S. Nosé, A Unified Formulation of the Constant Temperature Molecular Dynamics Methods, *J. Chem. Phys.*, **1984**, *81*, 511–519.

[223] W. G. Hoover, Canonical Dynamics: Equilibrium Phase-Space Distributions, *Phys. Rev. A*, **1985**, *31*, 1695.

[224] M. Parrinello and A. Rahman, Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method, *J. Appl. Phys.*, **1981**, *52*, 7182–7190.

[225] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, A Smooth Particle Mesh Ewald Method, *J. Chem. Phys.*, **1995**, *103*, 8577–8593.

[226] B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation, *J. Chem. Theory Comput.*, **2008**, *4*, 116–122.

[227] W. Humphrey, A. Dalke, and K. Schulten, VMD: Visual Molecular Dynamics, *J. Molec. Graphics*, **1996**, *14*, 33–38.

[228] J. E. Stone, An Efficient Library for Parallel Ray Tracing and Animation. Ph.D. Thesis, University of Missouri-Roula, 1998.

[229] M. Brehm and B. Kirchner, TRAVIS - A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories, *J. Chem. Inf. Model.*, **2011**, *51*, 2007–2023.

[230] B. Chen and J. I. Siepmann, Microscopic Structure and Solvation in Dry and Wet Octanol, *J. Phys. Chem. B*, **2006**, *110*, 3555–3563.

[231] J. N. Israelachvili, D. J. Mitchell, and B. W. Ninham, Theory of Self-Assembly of Hydrocarbon Amphiphiles into Micelles and Bilayers, *J. Chem. Soc. Faraday Trans. 2*, **1976**, *72*, 1525–1568.

[232] G. M. Grason and R. D. Kamien, Interfaces in Diblocks: A Study of Miktoarm Star Copolymers, *Macromolecules*, **2004**, *37*, 7371–7380.

[233] G. M. Grason, B. A. DiDonna, and R. D. Kamien, Geometric Theory of Diblock Copolymer Phases, *Phys. Rev. Lett.*, **2003**, *91*, 058304.

[234] S. T. Milner, Chain Architecture and Asymmetry in Copolymer Microphases, *Macromolecules*, **1994**, *27*, 2333–2335.

[235] H. Minehara, L. M. Pitet, S. Kim, R. H. Zha, E. W. Meijer, and C. J. Hawker, Branched Block Copolymers for Tuning of Morphology and Feature Size in Thin Film Nanolithography, *Macromolecules*, **2016**, *49*, 2318–2326.

[236] A. Dehghan and A.-C. Shi, Modeling Hydrogen Bonding in Diblock Copolymer/homopolymer Blends, *Macromolecules*, **2013**, *46*, 5796–5805.

[237] J. D. Vavasour and M. D. Whitmore, Self-Consistent Mean Field Theory of the Microphases of Diblock Copolymers, *Macromolecules*, **1992**, *25*, 5477–5486.

[238] F. S. Bates and G. H. Fredrickson, Block Copolymer Thermodynamics: Theory and Experiment, *Annu. Rev. Phys. Chem.*, **1990**, *41*, 525–557.

[239] A. K. Khandpur, S. Forster, F. S. Bates, I. W. Hamley, A. J. Ryan, W. Bras, K. Almdal, and K. Mortensen, Polyisoprene-Polystyrene Diblock Copolymer Phase-Diagram Near the Order-Disorder Transition, *Macromolecules*, **1995**, *28*, 8796–8806.

[240] M. W. Matsen and F. S. Bates, Unifying Weak- and Strong-Segregation Block Copolymer Theories, *Macromolecules*, **1996**, *29*, 1091–1098.

[241] V. Pryamitsyn and V. Ganesan, Self-Assembly of Rod–Coil Block Copolymers, *J. Chem. Phys.*, **2004**, *120*, 5824–5838.

[242] K. K. Tenneti, X. Chen, C. Y. Li, Y. Tu, X. Wan, Q. F. Zhou, I. Sics, and B. S. Hsiao, Perforated Layer Structures in Liquid Crystalline Rod–Coil Block Copolymers, *J. Am. Chem. Soc.*, **2005**, *127*, 15481–15490.

[243] M. W. Matsen and C. Barrett, Liquid-Crystalline Behavior of Rod-Coil Diblock Copolymers, *J. Chem. Phys.*, **1998**, *109*, 4108–4118.

[244] Y. Zhou, A. Nicolas, K. R. Thomas, and U. Steiner, Interplay of Electrohydrodynamic Structure Formation and Microphase Alignment in Lamellar Block Copolymers, *Soft Matter*, **2012**, *8*, 3841–3847.

[245] J. Y. Cheng, C. T. Rettner, D. P. Sanders, H. Kim, and W. D. Hinsberg, Dense Self-Assembly on Sparse Chemical Patterns: Rectifying and Multiplying Lithographic Patterns using Block Copolymers, *Adv. Mater.*, **2008**, *20*, 3155–3158.

[246] J. M. Kim, Y. Kim, W. I. Park, Y. H. Hur, J. W. Jeong, D. M. Sim, K. M. Baek, J. H. Lee, M. Kim, and Y. S. Jung, Eliminating the Trade-Off between the Throughput and Pattern Quality of Sub-15 nm Directed Self-Assembly via Warm Solvent Annealing, *Adv. Funct. Mater.*, **2015**, *25*, 306–315.

[247] M. P. Stoykovich, K. C. Daoulas, M. MuÌĹller, H. Kang, J. J. de Pablo, and P. F. Nealey, Remediation of Line Edge Roughness in Chemical Nanopatterns by the Directed Assembly of Overlying Block Copolymer Films, *Macromolecules*, **2010**, *43*, 2334–2342.

[248] M. W. Matsen and M. Schick, Stable and Unstable Phases of a Diblock Copolymer Melt, *Phys. Rev. Lett.*, **1994**, *72*, 2660.

[249] M. W. Matsen, Effect of Architecture on the Phase Behavior of AB-Type Block Copolymer Melts, *Macromolecules*, **2012**, *45*, 2161–2165.

[250] T. P. Russell, R. P. Hjelm Jr, and P. A. Seeger, Temperature Dependence of the Interaction Parameter of Polystyrene and Poly(methyl methacrylate), *Macromolecules*, **1990**, *23*, 890–893.

[251] J. Qin and D. C. Morse, Fluctuations in Symmetric Diblock Copolymers: Testing Theories Old and New, *Phys. Rev. Lett.*, **2012**, *108*, 238301.

[252] B. D. Olsen and R. A. Segalman, Self-Assembly of Rod–Coil Block Copolymers, *Mater. Sci. Eng., R*, **2008**, *62*, 37–66.

[253] K. Bolton, W. L. Hase, G. H. Peslherbe, and D. L. Thompson, *Modern Methods for Multidimensional Dynamics Computation in Chemistry;* World Scientific: New Jersy, 1998.

[254] Z. Yao, J. Wang, G. Liu, and M. Cheng, Improved Neighbor List Algorithm in Molecular Simulations using Cell Decomposition and Data Sorting Method, *Comput. Phys. Commun.*, **2004**, *161*, 27–35.

[255] Z. Cao, H. Wen, F. Wu, and D. H. C. Du; ALACC: Accelerating Restore Performance of Data Deduplication Systems Using Adaptive Look-Ahead Window Assisted Chunk Caching; In *16th USENIX Conference on File and Storage Technologies (FAST 18)*, pages 309–324, Oakland, CA, 2018. USENIX Association.

[256] L. Verlet, Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules, *Phys. Rev.*, **1967**, *159*, 98.

[257] B. Hess, C. Kutzner, D. Van Der Spoel, and E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *J. Chem. Theory Comput.*, **2008**, *4*, 435–447.

[258] A. A. Chialvo and P. G. Debenedetti, On the Use of the Verlet Neighbor List in Molecular Dynamics, *Comput. Phys. Commun.*, **1990**, *60*, 215–224.

[259] A. A. Chialvo and P. G. Debenedetti, On the Performance of an Automated Verlet Neighbor List Algorithm for Large Systems on a Vector Processor, *Comput. Phys. Commun.*, **1991**, *64*, 15–18.

[260] J. L. Finney, Long-Range Forces in Molecular Dynamics Calculations on Water, *J. Comput. Phys.*, **1978**, *28*, 92–102.

[261] W. F. van Gunsteren and H. J. Berendsen, Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry, *Angew. Chem., Int. Ed. Engl.*, **1990**, *29*, 992–1023.

[262] R. J. Petrella, I. Andricioaei, B. R. Brooks, and M. Karplus, An Improved Method for Nonbonded List Generation: Rapid Determination of Near-Neighbor Pairs, *J. Comput. Chem.*, **2003**, *24*, 222–231.

[263] V. Yip and R. Elber, Calculations of a List of Neighbors in Molecular Dynamics Simulations, *J. Comput. Chem.*, **1989**, *10*, 921–927.

[264] W. Mattson and B. M. Rice, Near-Neighbor Calculations Using a Modified Cell-Linked List Method, *Comput. Phys. Commun.*, **1999**, *119*, 135–148.

[265] T. N. Heinz and P. H. Hünenberger, A Fast Pairlist-Construction Algorithm for Molecular Simulations under Periodic Boundary Conditions, *J. Comput. Chem.*, **2004**, *25*, 1474–1486.

[266] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys.*, **1983**, *79*, 926–935.

[267] J. J. Morales and M. J. Nuevo, Comparison of Link-Cell and Neighbourhood Tables on a Range of Computers, *Comput. Phys. Commun.*, **1992**, *69*, 223–228.

[268] R. W. Hockney, S. P. Goel, and J. W. Eastwood, Quiet High-Resolution Computer Models of a Plasma, *J. Comput. Phys.*, **1974**, *14*, 148–158.

[269] D. C. Rapaport, Large-Scale Molecular Dynamics Simulation Using Vector and Parallel Computers, *Comput. Phys. Rep.*, **1988**, *9*, 1–53.

[270] B. Quentrec and C. Brot, New Method for Searching for Neighbors in Molecular Dynamics Computations, *J. Comput. Phys.*, **1973**, *13*, 430–432.

[271] M. Pütz and A. Kolb, Optimization Techniques for Parallel Molecular Dynamics Using Domain Decomposition, *Comput. Phys. Commun.*, **1998**, *113*, 145–167.

[272] R. Everaers and K. Kremer, A Fast Grid Search Algorithm for Molecular Dynamics Simulations with Short-Range Interactions, *Comput. Phys. Commun.*, **1994**, *81*, 19–55.

[273] A. Arnold and N. Mauser, An Efficient Method of Bookkeeping Next Neighbours in Molecular Dynamics Simulations, *Comput. Phys. Commun.*, **1990**, *59*, 267–275.

[274] W. F. Van Gunsteren and H. J. Berendsen, Computer Simulation as a Tool for Tracing the Conformational Differences between Proteins in Solution and in the Crystalline State, *J. Mol. Biol.*, **1984**, *176*, 559–564.

[275] J. L. Bentley, Multidimensional Binary Search Trees Used for Associative Searching, *Commun. ACM*, **1975**, *18*, 509–517.

[276] J. Barnes and P. Hut, A Hierarchical O(NlogN) Force-Calculation Algorithm, *Nature*, **1986**, *324*, 446–449.

[277] A. W. Appel, An Efficient Program for Many-Body Simulation, *SIAM J. Sci. and Stat. Comput.*, **1985**, *6*, 85–103.

[278] J. P. Gardner, A. Connolly, and C. McBride; A Framework for Analyzing Massive Astrophysical Datasets on a Distributed Grid; In *Astronomical Data Analysis Software and Systems XVI*, Vol. 376, page 69, 2007.

[279] Y. Kang, J. Nah, W. Park, and S. Yang, gkDtree: A Group-Based Parallel Update Kd-tree for Interactive Ray Tracing, *J. Syst. Archit.*, **2013**, *59*, 166–175.

[280] M. Shevtsov, A. Soupikov, and A. Kapustin; Highly Parallel Fast Kd-tree Construction for Interactive Ray Tracing of Dynamic Scenes; In *Computer Graphics Forum*, Vol. 26, pages 395–404. Wiley Online Library, 2007.

[281] N. Nakasato, Implementation of a Parallel Tree Method on a GPU, *J. Comput. Sci.*, **2012**, *3*, 132–141.

[282] C. A. R. Hoare, Algorithm 65: Find, *Commun. ACM*, **1961**, *4*(7), 321–322.

[283] B. Chen and J. I. Siepmann, A Novel Monte Carlo Algorithm for Simulating Strongly Associating Fluids: Applications to Water, Hydrogen Fluoride, and Acetic Acid, *J. Phys. Chem. B*, **2000**, *104*(36), 8725–8734.

[284] I. R. McDonald, NpT-Ensemble Monte Carlo Calculations for Binary Liquid Mixtures, *Mol. Phys.*, **1972**, *23*, 41–58.